

## **A Comparative Analysis of the Influence of High Stakes Testing Mandates in the Elementary School**

Albert Inserra and Kenneth R. Bossert

Albert Inserra, Professor, Dowling College, New York

Kenneth R. Bossert, Executive Director of Elementary Education, Three Village Central School District, Stony Brook, New York

### **Abstract**

The No Child Left Behind Act of 2001, sponsored by President George W. Bush, calls for 100 percent proficiency in reading and mathematics by 2014. This Federal mandate has caused all public schools in the United States to examine the programs in use to meet these requirements. In addition, states across the country have implemented a series of high-stakes testing to insure that school districts are accountable for measurable growth among the student population.

New York State has been on the front line in the implementation of high-stakes testing and increasing accountability. New York State is now administering assessments in English/Language Arts as well as Mathematics to all students in grades three through eight. This empirical research examined how elementary teacher attitudes towards high-stakes testing influence the range of instructional methods they employ and their self-reported ability to engage students in enrichment activities within the classroom.

### **Background/Introduction**

In 1983, The National Commission on Education published a document titled *A Nation at Risk*, which raised public awareness regarding a “crisis” taking place within American public schools. American students were viewed as falling behind students attending public schools in other countries, placing American students in a position to lose economic and scientific opportunities to students in countries such as Japan and China.

Eighteen years later, similar concerns still exist as American students struggle to compete within the global marketplace. In response, President George W. Bush proposed federal legislation known as the No Child Left Behind Act, passed by Congress in 2001. This act mandated that all public school students would be proficient in mathematics and reading by 2014 as evidenced by standardized, high-stakes testing that would be implemented at designated grade levels. State education departments across the country responded and developed assessments that were to be administered in the areas of English/Language Arts and Mathematics across selected grade levels.

In January of 2001, newly inaugurated President George W. Bush announced *No Child Left Behind* as the “cornerstone of my Administration.” “Despite nearly \$200 billion in Federal spending since the passage of the Elementary and Secondary Education Act of 1965 (ESEA), too many of our neediest children are being left behind” (United States Education Dept., 2004, [www.ed.gov/print/nclb/overview/intro/execsumm.html](http://www.ed.gov/print/nclb/overview/intro/execsumm.html), para. 2).

The NCLB Act:

Reauthorizes the ESEA, incorporates the principles and strategies proposed by President Bush. These include increased accountability for States, school districts, and schools; greater choice for parents and students, particularly those attending low-performing schools, more flexibility for state and local educational

agencies (LEAs) in the use of federal education dollars; and a stronger emphasis on reading, especially for our youngest children” (United States Education Dept., 2004, [www.ed.gov/print/nclb/overview/intro/execsumm.html](http://www.ed.gov/print/nclb/overview/intro/execsumm.html), para. 6).

In 2005, the United States Education Department, as a function of NCLB, reviewed student progress in each state since the passage of NCLB. The United States Education Department noted the following gains that took place between 2002 and 2004.

Fourth-grade mathematics achievement increased by 11 percentage points; the black-white achievement gap in fourth-grade reading narrowed by three percentage points; the Hispanic-white achievement gap in fourth-grade reading narrowed by five percentage points; the black-white achievement gap in fourth-grade mathematics narrowed by 10 percentage points; and the Hispanic-white achievement gap in fourth-grade mathematics narrowed by 10 percentage points (United States Education Dept., 2004, [www.ed.gov/print/nclb/overview/intro/execsumm.html](http://www.ed.gov/print/nclb/overview/intro/execsumm.html), para.14).

High-stakes tests have been administered in New York State since 1997, when Commissioner Richard Mills designated the fourth and eighth grades to be tested in both English/Language Arts and Mathematics. These tests were used in addition to existing mandated Regents testing to meet graduation requirements. The mere implementation of these uniform assessments, according to Mills, increased school district accountability and public awareness regarding school performance. For the first time, parents and community members would have a concrete diagnostic tool to measure how their local school district compared with similar schools in neighboring districts, and across the state. The “stakes” were raised to an even higher level of accountability when local newspapers began to publish New York State report cards in Sunday editions annually each spring.

In January of 2006, the number of students that would be subjected to standardized New York State assessments grew tremendously. Testing would no longer focus solely on fourth and eighth graders, and would include all students in grades three through eight. January was the month designated for a New York State assessment in English/Language Arts in grades three through eight, while March would feature a New York State assessment in Mathematics in those same grades.

Students who failed to meet minimum New York State expectations (Levels 3 or 4 - meeting or exceeding expectations) would be mandated to receive academic intervention services. Schools whose test scores failed to demonstrate “adequate yearly progress” (AYP, as defined by the New York State Education Department as acceptable progress by a district/school toward the goal of proficiency for all students) would be labeled as “schools needing improvement,” a designation that would not please administration, teachers, or the community.

The potential for individual teacher accountability thus became possible in classrooms. The data that were generated from students’ test scores could be evaluated on a “teacher by teacher” basis. Administration could easily disaggregate test results to analyze which class scores met expectations and which did not. These testing results could influence teacher observations, evaluation, and recommendations for tenure.

For this study, the Teacher Evaluation Instrument published by Charlotte Danielson in her book, *Enhancing Professional Practice: A Framework for Teaching* (1996) was used as a “measurement” for good teaching practices. Danielson breaks down effective teaching into four

domains (planning and preparation, classroom environment, instruction, and professional responsibilities). Within each domain, many components and elements are described. These components and domains describe expected teaching practices in each of the four key areas. Teachers can be rated as unsatisfactory, basic, proficient, or distinguished using the Danielson language.

Danielson is a former economist and an educational consultant. She is a consultant and guest speaker with engagements scheduled throughout the United States on a consistent basis. She has taught at all levels, from kindergarten through college, and has worked as an administrator, a curriculum director, and a staff developer. In her consulting work, Danielson has specialized in aspects of teacher quality and evaluation, curriculum planning, performance assessment, and professional development. Danielson is the author of a number of books supporting teachers and administrators. Danielson's *Standards of Teaching* (1996) have been adapted by many districts throughout the United States to be used as a teacher evaluation instrument. Danielson's *Standards of Teaching* (1996) were used to develop the survey instrument for this study.

### **Theoretical Underpinnings**

In November of 2006, the Center on Education Policy published a report titled: *Ten Big Effects of the No Child Left Behind Act on Public Schools*. Written five full years after the passing of this legislation, this report served as a comprehensive assessment of what effect the Act has had on public education in the United States.

Jack Jennings, the President of the Center on Education Policy, authored the study which appeared in Phi Delta Kappan magazine in October of 2006. Jennings (2006) states: "One; State and district officials report that student achievement on state tests is rising, which is cause for optimism. It's not clear, however, that students are really gaining as much as rising percentages of proficient scores would suggest" (Jennings, 2006, p. 1). Jennings (2006) explains that local districts credit their own policies and changes in practice for the increase in measurable student achievement.

Two; schools are spending more time on reading and math, sometimes at the expense of subjects not tested...71 percent of districts are reducing time spent on other subjects in elementary schools – at least to some degree. The subject most effected is social studies, while physical education is least effected" (p. 1).

Jennings (2006) indicates that 60 percent of districts require a specific amount of time for reading in the elementary schools.

"Three; schools are paying much more attention to the alignment of curriculum and instruction and are analyzing test score data much more closely" (p. 2). The most common improvement, according to Jennings, is greater alignment of curriculum and instruction with standards and assessments.

Four; low-performing schools are undergoing makeovers rather than the most radical kinds of restructuring. Five; schools and teachers have made considerable progress in demonstrating that teachers meet the law's academic qualifications – but many educators are skeptical this will really improve the quality of teaching (p.2).

"Six; students are taking a lot more tests. In 2002, 19 states had annual reading and mathematics tests in grades three through eight and once in high school; by 2006, every state had

such testing” (p. 3). “Seven; schools are paying much more attention to achievement gaps and learning needs of particular groups of students...Eight; the percentage of schools on state ‘needs improvement’ lists has been steady but it not growing” (p. 4).

Jennings continues by examining the role of the federal government, claiming: “Nine; the federal government is playing a bigger role in education. As a result of NCLB, the federal government is taking a much more active role in public elementary and secondary education than in the past” (p. 5).

Ten; NCLB requirements have meant that state governments and school districts also have expanded roles in school operations, but often without adequate federal funds to carry out their duties. State governments are also taking a much more active role in public education, because they must carry out NCLB provisions that affect all their public schools” (p. 5).

These responsibilities now thrust upon the state include: creating or expanding three through eight testing programs, setting minimum testing goals, providing assistance to students, groups, and schools in need, and establishing criteria to determine whether current teachers meet NCLB’s teacher-quality requirements. Jennings (2006) indicates that local school districts must also assume more duties than before NCLB.

In March of 2006, The Center on Education Policy (CEP) released *From the Capital to the Classroom: Year 4 of the No Child Left Behind Act*. This became a comprehensive, long-term national study of the Act.

In order to reach conclusions regarding NCLB, CEP used the following research methods: “a survey of all 50 states, a nationally representative survey of 299 school districts, case studies of 38 geographically diverse districts and 42 schools, six special analyses of critical issues in implementing the Act, and three national forums” (Stark & Jennings, 2006, p. 8).

Stark and Jennings (2006) state that four broad conclusions can be made following their national study:

The impact of the *No Child Left Behind Act* continued to widen and deepen during 2005, the law’s fourth year of implementation. NCLB affects a range of state and local decisions, both small and large – when and how students take tests, which textbook series districts adopt, which children receive extra attention and how they are grouped, how states and districts spend their money, how teachers are trained, and where principals and teachers are assigned to work, to cite just some examples (Stark & Jennings, 2006, p. 12).

The first conclusion of the CEP to be examined is the impact NCLB has had on curriculum and instruction taking place within public schools across the nation. The CEP clearly states that teaching and learning have changed as a result of NCLB. Teachers and administrators have increased the time and effort spent on curriculum alignment with state standards and assessments. Data based instructional planning is much more prevalent than prior to NCLB. A narrower curriculum has been created. “Seventy-one percent of the school districts surveyed reported that they have reduced elementary school instructional time in at least one other subject to make more time for reading and mathematics – the subjects tested for NCLB” (Stark & Jennings, 2006, p. 21). In some struggling districts, the case study analysis indicates that students receive double periods of math or E. L. A. instruction.

The CEP report notes that teachers are being given greater direction. Pacing guides, instructional coaches, and mentor teachers have become more prevalent. The NCLB Act has also influenced what teachers must do to be considered well-qualified. By the Act's definition, soon all academic subject teachers must meet criteria to be considered highly qualified. This includes certification (or dual certification) in the area being instructed, ongoing professional development, and mentoring.

The second broad conclusion made by the Center on Education Policy (CEP) assesses student achievement.

Scores on state tests have risen in a large majority of states and school districts, according to the state and local officials surveyed. Many survey respondents cited the NCLB requirements for adequate yearly progress (AYP) as an important factor in rising achievement, but far more credited school district policies and programs as important contributors to these gains (Stark & Jennings, 2006, p. 31).

This second conclusion of the report associates increases in achievement scores and more students falling into the proficient level or above as evidence that learning has increased.

The CEP study concluded that the effects of NCLB are "holding steady." Conclusion three states: "The number of schools identified for improvement under the NCLB accountability provisions has remained fairly steady since last year, despite earlier predictions that these numbers would soar over time. Participation rates in school choice and tutoring remain low" (Stark & Jennings, 2006, p. 32).

The CEP states that "the schools identified for improvement are not always the same schools each year. A modest proportion of schools tests out of improvement each year, while other new schools enter improvement categories. Overall, the percentage and number of schools in improvement have varied little" (Stark & Jennings, 2006, p. 33). The scenario can create a confusing statistical report. The percentage of schools indicated as needing improvement may remain constant, while many schools are making adequate yearly progress (AYP).

The CEP reports that there has been little change in choice and tutoring participation.

The percentage of all eligible students taking advantage of the NCLB school choice option to change schools remains at less than 2 percent, while the percentage participating in supplemental educational services (tutoring) has hovered around 20 percent for the past two years (Stark & Jennings, 2006, p. 40).

The fourth major conclusion to be drawn from the CEP study regards the greater impact NCLB has had on urban school districts. The majority (54 percent) of Title I schools identified for improvement nationwide are located in urban districts:

A disproportionate share because only 27 percent of Title I schools are located in urban districts. Greater proportions of urban districts than suburban or rural districts have been identified for district improvement. About 90 percent of the schools in restructuring, the last status category of NCLB's sanctions, are in urban districts" (Stark & Jennings, 2006, p. 44).

Diversity found in urban districts, according to the CEP, is a major reason why NCLB is having a greater impact. Meaning the accountability set forth by NCLB is leading to changes in school policy. Urban districts are also more affected by NCLB sanctions because of their size.

They must demonstrate adequate yearly progress (AYP) for dozens of schools, while smaller districts may have only one school.

In addition to the four broad conclusions drawn by the CEP, the study also arrived at several major findings about specific aspects of NCLB, some are interpreted as positive, while others are viewed (by the CEP) as negative. Positive effects include high expectations of students and subgroups within the student population. The performance of subgroups that had traditionally “lagged behind” (e.g., special education students) resulted in a greater focus on subgroup performance. The study indicated improved alignment and use of test data.

The CEP draws three major conclusions that are classified as “negative.” The first refers to a “greater burden, without adequate funding.” The CEP states:

States and districts lack both the funding and the staff capacity to carry out all of the demands of NCLB, according to our surveys. Some 80 percent of school districts said they had costs for NCLB that were not covered by federal funds. Thirty-three states reported that federal funds have been inadequate to assist all schools identified for improvement, and less than half of school districts said they have enough money to assist identified schools at least somewhat. In addition, 36 states told us they do not have enough staff to implement NCLB – a major concern because state agencies are the source that school districts most often turn to for help in implementing NCLB. (Stark & Jennings, 2006, p. 44)

Teacher stress and staff morale were included in other conclusions drawn by CEP. Teachers who were interviewed noted that constant pressure to raise test scores and the negative connotation that comes along with being labeled a “SINI” (School in Need of Improvement) has a negative effect on morale in schools.

CEP cited challenges with accountability. Specifically, survey respondents were concerned about how progress was judged for students with disabilities and English language learners. A major concern among these respondents was the ability to bring 100 percent of students to proficiency by 2014, as NCLB dictated.

Through the survey of 299 schools throughout the 50 states, the Center of Education Policy provided data to show many trends. These included:

78 percent reported an increase in student achievement from 2003-2004 to 2004-2005. Thirty-five states reported gains in reading, while 36 states reported gains in Mathematics. Twenty-four percent of all school districts did not make adequate yearly progress based on 2004-2005 testing. Just 1.6 percent of the students who were eligible for NCLB choice in 2005-2006 actually took advantage of it (Stark & Jennings, 2006, p. 55).

The CEP made eight key recommendations for change in administration and funding for NCLB mandates. Recommendation One: Transparency in state accountability plans. Two: Monitoring effects of flexibility on AYP. Three: Guidance and support for modified standards and assessments. Four: The President and the Congress must provide adequate funding for the Act. Five: Support (from Congress and the Education Department) for school improvement. Six: The Department and Congress should give states and school districts sufficient resources and authority to successfully oversee supplemental educational service providers and evaluate their effectiveness in raising student achievement. Seven: Supplemental services pilot program. Eight: Attention to other subjects (Stark & Jennings, 2006, p. 58).

The study conducted by the Center on Education Policy (CEP) is perhaps the most comprehensive evaluation of the widespread effect that the No Child Left Behind Act has had on a nationwide basis after four years of full implementation. The CEP relied heavily upon the assumption that the high-stakes testing in place was a reliable resource to measure student achievement and school growth.

### **Empirical Research**

This study examined how teacher attitudes towards the No Child Left Behind Act and high-stakes testing relate to select teaching practices that include: planning, preparation, classroom environment, and instruction. Teachers in both high-stakes testing environments and those in classrooms that do not have a high-stakes test were studied. Further, this study examined the relationships among teacher attitudes toward No Child Left Behind accountability and their practices in planning, preparation, classroom environment, instruction, test preparation activities and enrichment practices. This: empirical research was focused on two major research questions:

How do teachers in high-stakes testing environments and teachers in a non high-stakes testing environment differ in practices they employ regularly for instruction, enrichment behaviors, test preparation activities, classroom environment, attitudes toward NCLB/high-stakes testing, planning, and preparation?

In a both a high-stakes and non high stakes testing environment, are there relationships among teachers' years of experience, tenure status, instruction, enrichment behaviors, test preparation activities, classroom environment, attitudes toward NCLB/high-stakes testing, planning, and preparation?

Subjects were chosen in a purposeful sample from teachers in a high-stakes testing environment and those in an environment that does not have a high-high stakes test, from a public school district currently engaged in the administration of New York State assessments in English/Language Arts and Mathematics, as well as one that evaluates teachers by employing the Charlotte Danielson framework for effective teaching in the formal supervision process within the school district.

Teachers in the selected district have received professional development in Danielson's *Standards of Teaching* for the school years 2004-2005 through the present, including each of the four domains and rubric instrument used for evaluation. . Administration in the district uses Danielson's language to evaluate professional staff.

The school district is comprised of three elementary schools that were studied: a K-2 primary school, a 3-6 intermediate school, and a K-6 elementary school. The total number of potential respondents was 152. The actual number of respondents was 143. This population included only general classroom teachers in grades kindergarten through the sixth grade, including special education teachers and their certified teacher assistants.

### **Data Analysis**

“How do teachers in high-stakes testing environments and teachers in a non high-stakes testing environment differ in practices they employ regularly for instruction, enrichment behaviors, test preparation activities, classroom environment, attitudes toward NCLB/high-stakes testing, planning, and preparation?” presented the most interesting and relevant findings.

This research question sought to identify differences in practices from those respondents (teachers) in high-stakes testing environments (grades 3 through 6) from those in non high-stakes testing environments (grades K-2).

An independent sample *t-test* was conducted in order to draw conclusions about the differences between the two groups (high-stakes and non high-stakes). Table 1 indicates a mean comparison between respondents in the high-stakes and non high-stakes groups. As illustrated in Table 1, there was no significant difference between respondent groups (high-stakes and non high-stakes) on the factors of: instruction, classroom environment, attitude towards NCLB/high-stakes testing, planning, and preparation.

The two factors where a significant difference was identified are the enrichment and test preparation factors.

**Table 1**

*Independent Sample T-Test – Subgroup Comparison of High-Stakes and Non- High-Stakes Teachers*

	N	Mean	Standard Deviation	<i>t-test</i>	Significance
Instruction					
Non high-Stakes	54	66.00	6.49	-.665	.507
High-Stakes	86	66.69	5.57		
Enrichment					
Non high-Stakes	54	49.04	5.81	5.17	.000
High-Stakes	78	42.36	8.17		
Test Preparation					
Non high-Stakes	54	21.61	5.48	-11.4	.000
High-Stakes	85	31.47	4.61		
Environment					
Non high-Stakes	57	37.63	4.18	1.23	.221
High-Stakes	85	36.73	4.35		
NCLB/H.S.T.					
Non high-Stakes	49	25.39	6.21	-.167	.867
High-Stakes	75	25.56	5.17		
Planning					
Non high-Stakes	56	22.46	2.17	-.293	.770
High-Stakes	83	22.58	2.30		
Preparation					
Non high-Stakes	57	21.74	2.34	-1.24	.217
High-Stakes	82	22.22	2.19		

Table 1 indicates that respondents in each subgroup (non high-stakes and high-stakes) responded to items within the five factors of instruction, environment, NCLB/high-stakes testing,



planning, and preparation in very similar fashion. By comparing means this is evident. All respondents “slightly agree” with the issues contained in the items and also agree with each other.

With significance levels of .000, it is clear that teachers in each group (non high-stakes and high-stakes) differed in the way they responded to items within the enrichment and test preparation factors.

In the enrichment factor, with a mean score of 49.04 (range of 13 to 65), the non high stakes teachers “agreed” more with the items in this variable than their counterparts in the high-stakes group, who had a mean score of 42.36.

In the test preparation factor, the non high-stakes teachers slightly agreed with the items in this factor with a mean score of 21.61 (with a range of 9 to 45), while the high-stakes testing respondents had a mean score of 31.47, showing an agreement with the items in this factor.

In order to better understand the teacher responses, an item analysis was conducted on the two variables that showed a significant difference.

Within the variable of enrichment, nine items were all shown to be significant at the .05 level. Tables 1.2 and 1.3 illustrate the items found to be significant within the enrichment and test preparation variables.

**Table 1.2** *T-test for equality of means/group statistics, Enrichment Activities*

Item #	Item Language	Non-High Stakes Mean	High-Stakes Mean	Mean Diff.	<i>t-test</i>	Sig. (2-tailed)
6	I plan with an interest in the cultural heritage of each student.	3.96	3.60	0.36	2.23	.028
67	My class uses cooperative learning strategies regularly	4.31	4.01	0.30	2.19	.030
68	My class engages in dramatic play activities (Reader's theater, etc.).	3.84	2.52	1.32	7.51	.000
70	My class has the opportunity to engage in enrichment activities that may enrich their civic knowledge.	4.19	3.34	0.85	5.14	.000
71	My class seems to enjoy the instructional activities that are part of our normal routine	4.40	3.97	0.43	3.21	.002
73	My class is provided with instruction that can be considered thematic.	4.18	3.60	0.58	3.70	.000
74	My class has parents / guardians that play an important role.	4.28	3.41	0.87	5.27	.000
75	My class has opportunities to be engaged in music based activities.	3.89	3.07	0.82	4.83	.000
76	My class has opportunities to be engaged in art activities.	4.19	3.47	0.72	4.82	.000

Within these nine items, respondents within the non high-stakes group had a significantly higher mean score for enrichment activities than the high-stakes respondents.

**Table 1.3**  
*T-test for Equality of Means / Group Statistics*  
*Test Preparation*

Item #	Item Language	Non high Stakes Mean	High-Stakes Mean	Mean Diff.	t-test	Sig. (2-tailed)
55	My class spends a large portion of the day engaged in test prep. activities	2.92	3.73	0.81	-4.24	.000
56	My class stresses the importance of high achievement on NYS Assessments	2.55	3.39	0.84	-4.58	.000
57	My class uses materials that are directly aligned with NYS Assessments	3.10	4.35	1.25	-7.41	.000
60	My class receives an appropriate amount of test preparation materials	2.82	4.10	1.28	-7.90	.000
61	My class does not engage in test preparation activities	2.70	1.51	1.19	6.68	.000
62	My class has homework aligned to NYS assessments	2.49	2.90	0.41	-2.10	.042
63	My class frequently discusses the NYS Assessments	1.64	3.31	1.67	-9.71	.000
64	My class is well aware of the content and format of the NYS assessment in ELA	1.73	4.22	2.49	-16.72	.000
65	My class is well aware of the content and format of the NYS assessment in Mathematics	1.70	3.93	2.23	-13.77	.000

As Table 1.3 shows, within these nine items, the high-stakes group had a significantly higher mean score for test preparation activities than the non high-stakes group of respondents.

It is clear that the items included in the enrichment and test preparation variables require careful consideration when drawing conclusions about this study.

A frequency analysis was conducted in order to more closely examine the responses from each group studied. The distribution of responses, from “strongly disagree” to “strongly agree,” was examined.

Within the factors of enrichment activities and test preparation activities, there were several items that required further investigation into the frequency of distribution of responses offered by non high-stakes (primary) and high-stakes (intermediate) teachers.

Within the Enrichment Activities factor, Item 68: “My class engages in dramatic play activities,” 63.1 percent of the primary teachers agreed to strongly agreed with this item, while 20.9 percent of intermediate teachers believed their students engaged in dramatic play activities.

Within item 70: “My class has the opportunity to engage in enrichment activities that may enrich their civic knowledge,” 85.9 percent of primary teachers agreed to strongly agreed, while 53.5 percent of intermediate teachers agreed to strongly agreed with this item.

Within item 71: “My class seems to enjoy the instructional activities that are part of our normal routine,” 91.2 percent of primary teachers agreed to strongly agreed with this item, while 79.1 percent of intermediate teachers agreed to strongly disagreed with this item, indicating that students do not enjoy the instructional activities within their intermediate classroom as much as the primary teachers.

Within item 73: “My class is provided with instruction that can be considered thematic,” 79 percent of primary teachers agreed to strongly agreed with this item, while 62.8 percent of intermediate teachers agreed to strongly agreed with this item.

Within item 74: “My class has parents/guardians that play an important role,” 82.4 percent of primary teachers agreed to strongly agreed with this item, while 46.5 percent of intermediate teachers agreed to strongly agreed that parents/guardians play an important role, indicating their belief that parents play a more important role at the primary level.

Within item 76: “My class has opportunities to be engaged in art activities,” 82.5 percent of primary teachers agreed to strongly agreed, while only 50 percent of intermediate teachers agreed to strongly agreed that their class has opportunities to engage in art activities.

As Table 1.3 indicated there are significant differences in the way teachers in a non high stakes testing environment and teachers in a high stakes testing environment responded to items within the Test Preparation Activities factor.

Within item 55: “My class spends a large portion of the day engaged in test preparation activities,” 29.8 percent of primary teachers agreed to strongly agreed with this item, while 66.2 percent of intermediate teachers agreed to strongly agreed that a large portion of their instructional day is devoted to test preparation activities.

Within item 56: “My class stresses the importance of high achievement on NYS assessments,” 17.5 percent of primary teachers agreed to strongly agreed with this item, while 51.2 percent of intermediate teachers agreed to strongly agreed with this item, indicating intermediate teachers agreed that they did stress the importance of high achievement on NYS assessments. If slightly agree were added, 90.6 percent of teachers would tend to agree with this item at the intermediate level.

Within item 57: “My class uses materials that are directly aligned with NYS assessments,” 40.3 percent of primary teachers agreed to strongly agreed with this item, while 90.7 percent of the intermediate teachers agreed or strongly agreed. This demonstrates that intermediate teachers agree that the instructional items they use in class are directly aligned with NYS assessments.

Within item 60: “My class receives an appropriate amount of test preparation materials,” 20.6 percent of primary teachers agreed to strongly agreed with the item, while 82.6 percent of intermediate teachers agreed or strongly agreed that their class receives an appropriate amount of

preparation materials. The way respondents defined “appropriate” may have influenced responses to this item.

Within item 63: “My class frequently discusses the NYS assessments”, 1.8 percent of primary teachers agreed to strongly agreed with this item, while 43 percent of intermediate teachers agreed to strongly agreed with this item. This demonstrates that primary teachers tend not to discuss the NYS assessments with their students.

The second research question : In both a high stakes and non high stakes testing environment, are there relationships among teachers’ years of experience, tenure status, instruction, enrichment behaviors, test preparation activities, classroom environment, attitudes toward NCLB/high stakes testing, planning and preparation?

Important information can be determined by comparing the two groups of respondents (high-stakes and non high-stakes) correlations to enrichment for each factor as illustrated in Tables 1.4 and 1.5. The Tables that follow illustrate comparisons that can be drawn by comparing the two groups of respondents; high-stakes and non high-stakes teachers.

**Table 1.4**  
*Correlations; Comparing Respondent Groups*  
*High-Stakes and Non high-Stakes Respondents*

Group	Instruction	Test- Prep.	Classroom Environment	Attitude NCLB	Planning	Preparation
Enrichment High- Stakes	.356	.072	.214	.527	.351	.219
Enrichment Non high- Stakes	.698	.166	.445	.362	.487	.407

**Table 1.5**  
*Correlations; Comparing Respondent Groups*  
*High-Stakes and Non high-Stakes Respondents*  
*Percent of Variance for Enrichment*

Group	Instruction	Test- Prep.	Classroom Environment	Attitude NCLB	Planning	Preparation
Enrichment High- Stakes	13%	1%	5%	28%	12%	5%
Enrichment Non high- Stakes	49%	3%	20%	13%	24%	17%

Many conclusions can be drawn by comparing the respondents from each group (high-stakes and non high-stakes testing).

Adhering to Danielson's instructional practices accounts for 49 percent of the variance in enrichment activities in the non high-stakes setting and only 13 percent of the variance in enrichment activities in the high-stakes testing setting. Overall, this would indicate that a non high-stakes teacher is able to use Danielson's instructional strategies to implement enrichment activities more consistently than a high-stakes teacher.

Adhering to Danielson's definition of a model classroom environment accounts for 20 percent of the variance in enrichment activities in the non high-stakes setting, and only five percent of the variance in enrichment activities in the high-stakes testing setting.

A teacher's attitude towards the No Child Left Behind Act accounts for 28 percent of the variance in enrichment activities in the high-stakes testing environment, and only 13 percent of the variance in enrichment activities in the non high-stakes testing environment. This indicates that teachers' attitudes towards NCLB are correlated with their ability to conduct enrichment activities within the classroom.

Adhering to Danielson's definition of planning accounts for 24 percent of the variance in enrichment activities in the non high-stakes setting, and only 12 percent of the variance in enrichment activities in the high-stakes testing setting.

Adhering to Danielson's definition of preparation accounts for 17 percent of the variance in enrichment activities in the non high-stakes setting, and only five percent of the variance in enrichment activities in the high-stakes testing setting.

Teachers in the non high-stakes and high-stakes respondent groups did not show a positive correlation between enrichment activities and test preparation activities.

Among non high-stakes teachers, Danielson's instructional practices, classroom environment, planning and preparation are all strongly related to enrichment activities. The higher percentage of the variance accounting for enrichment within the non high-stakes testing group as compared to teachers in the high-stakes testing group indicates an ability to conduct enrichment activities within a non high-stakes testing classroom while practicing Danielson's teaching strategies.

Among high-stakes teachers, a positive attitude towards NCLB has a strong relationship with enrichment, and there is also a small positive relationship for Danielson's instructional practices and enrichment.

## **Findings**

This study attempted to ascertain relationships between a teacher's placement within a high-stakes testing environment, third through the sixth grade, or a non high-stakes placement, grades kindergarten through second grade and seven distinct variables: instruction, enrichment activities, test preparation activities, classroom environment, attitude towards No Child Left Behind/high-stakes testing, planning, and preparation.

Teachers in a high-stakes testing environment and teachers in a non high-stakes testing environment were asked if they differ in practices they employ regularly for planning, preparation, classroom environment, instructional practices, test preparation activities, enrichment behaviors, and their attitudes toward NCLB/high-stakes testing. Survey responses indicated that there is no significant difference in practices employed within planning, preparation, classroom environment, and attitudes toward NCLB/high-stakes testing. There were, however, significant differences between non high-stakes and high-stakes testing teachers' responses in the areas of test preparation activities and enrichment activities. The primary

teachers who did not work in a high-stakes teaching environment reported fewer activities in test preparation and more enrichment activities in their classes

### **Conclusions and Recommendations**

It is difficult for school administrators to effectively measure what is being “lost” due to the amount of instructional time that is dedicated to subjects that are assessed by a high-stakes test. This study indicated that within a high-stakes-testing environment, teachers focus the largest percentage of instructional time on the subjects that are assessed by the State of New York, namely, English/Language Arts and Mathematics. Stark et al. (2006) suggest this loss may come as a result of decreased instructional time in Science, Social Studies, and enrichment activities.

Comparing non high-stakes teachers’ attitudes to high-stakes teachers’ attitudes in the area of NCLB/high-stakes testing, minimal differences are observed. It is clear that self-reporting of positive attitudes towards Danielson’s dimensions of planning, preparation, classroom environment and instruction are related to a positive self-reporting of enrichment activities by non high-stakes teachers.

. High-stakes teachers reported that their classes spend a larger portion of the day engaged in test preparation activities, and stress the importance of high achievement on the New York State assessments more than non high-stakes teachers. In addition, high-stakes teachers frequently use more materials that are directly aligned with New York State assessments, their classes receive a larger amount of test preparation materials, their assigned homework is more directly aligned with New York State assessments, their classes more frequently discuss New York State assessments, and their students are much more aware of the content of New York State assessments in both English/Language Arts and Mathematics.

The following recommendations can be proposed based upon the findings and conclusions of this study conducted in a suburban, Long Island, New York school district at the Kindergarten through sixth grade levels:

1. There are clear differences in the amount of enrichment activities that teachers are able to offer at the primary and intermediate levels. Teachers who instruct within a high-stakes (testing) environment need to make a conscious effort to focus on areas of the curriculum that include a variety of enrichment activities.

2. Although the district studied has offered professional development using Danielson’s model, additional training for teachers is recommended. It is also recommended that districts that do not currently offer professional development for the professional staff using Danielson’s Standards of Teaching (1996) consider making this a part of a balanced professional development plan to target appropriate enrichment activities. Danielson (1996) suggests that one of the key uses of her Framework for Teaching is structuring professional development opportunities. The original survey used for this study supports that notion that teachers can identify areas within their own teaching skills that should improve when they have a well developed framework in which they have been trained. Other educational theorists and their studies should be considered alongside of Danielson as well.

3. The results of the surveys should be shared with elementary principals, middle school principals, and central office administrators in public school districts across the State of New York and beyond to help plan future professional development opportunities. The surveys helped to identify issues that teachers felt were important. Effective professional development programs could be planned based on the data teachers provided in any district or school.

4. The results of this study should be shared with decision makers in the field of education so that they may see the effects high-stakes testing have on teacher attitudes and instructional practices at the elementary level. If the goal of an elementary school program is to provide students with an enriching learning opportunity, high-stakes testing is affecting this mission.

5. School administrators should investigate methods to prepare students for the New York State assessments within the areas of Science, Social Studies, Music, and Art. This may help to provide more enrichment opportunities to students without such a large focus on English/Language Arts and Mathematics. Instruction in the subject areas that are assessed should not be limited to a set portion of the instructional day.

6. The non high-stakes teachers surveyed seemed to be more involved in enrichment activities more so than the teachers at the intermediate level. Non high-stakes level teachers should have an opportunity to share strategies with their colleagues at the intermediate level. Strategies more frequently used at the primary level (such as learning centers, thematic teaching, and cooperative learning) should be employed with greater frequency at the intermediate level.

## REFERENCES

- Center on Education Policy (2006). From the capital to the classroom. Year 4 of the No Child Left Behind Act. March, 2006. Retrieved June 2, 2007, from the EBSCOhost database.
- Chudowsky, N., Chudowsky V. & Kober, N. (2007). From the capital to the classroom. Year 5 of the No Child Left Behind Act. July, 2007. Retrieved August 30, 2007, from the EBSCOhost database.
- Cochran-Smith, M. & Lytle, S. (2006). Troubling images of teaching in No Child Left Behind. *Harvard Educational Review*, 76, 669-697
- Danielson, C. (1996). Teacher Evaluation Instrument: Enhancing professional practice: a framework for teaching. Retrieved January 11, 2007, from EBSCOhost database.
- Goldberg, M. (2006). Losing students to high-stakes testing. *Education Digest*, 70 (7) 10-19.
- Greene, Winter, & Foster. (2003). Testing high stakes tests: can we believe the results of accountability tests? Center for Civic Innovation at the Manhattan Institute. February, 2003. Retrieved May 24, 2007 from EBSCOhost database.
- Holland, P. (2002). Two measures of change in the gaps between the CDFs of test-score distributions. *Journal of Educational and Behavioral Statistics*, 27(1), 3-17.
- Jennings, J. (2006). Ten Big Effects of the No Child Left Behind Act on Public Schools. *Phi Delta Kappan*, October 2006, 1-5.
- Jerald, C. (2006, July). Teach to the Test? Just say no: The Center for Comprehensive School Reform and Improvement, 1-6.
- Kantor, H. and Lowe, R. (Winter, 2006). From new deal to no deal; No Child Left Behind and the devolution of responsibility for equal opportunity. *Harvard Educational Review*, 26, 474-502.
- Kinthead, C. (2004). No Child Left Behind: The Oxymoron of Accountability. *Educational Equity*, 1, 1-8.
- Koger, L., Campbell, H., and Thacker, A. (2007) Behind the numbers: interviews in 22 states about achievement data and No Child Left Behind Act policies. *Human Resources Research Organization*, Alexandria, VA, 1-31.
- Koretz, Daniel (2005). Alignment, High Stakes, and the Inflammation of Test Scores. *National Center for Research on Evaluation*, 655, 1-22.
- McMillan, J. (2005). The impact of high-stakes test results on teachers' instructional and classroom assessment practices. *Metropolitan Educational Research Consortium*, 1-26.
- McMurrer, J. (2007). Choices, Changes, and Challenges: Curriculum and Instruction in the NCLB Era. Center of Education Policy. Washington D.C. Retrieved August 30, 2007 from the EBSCOhost database.



- Montgomery, P., Ranney, L., & Growe, R. (2003). High-stakes tests versus high-quality education. *University of Louisiana at Lafayette*. Retrieved November 17, 2006, from ERIC database.
- O'Neil, John and Tell, Carol (1999). Why students lose when "tougher standards" win. *Educational Leadership*, ASCD, September, 1999, 18-22.
- Senge, P., Lannon-Kim, C. (1991). Recapturing the spirit of learning through a system approach. *School Administrator*, 48 (9), 8-13.
- Smith, M.L., Edelsky, C., Draper, K., Rottenberg, C., & Cherland, M. (1989). The role of testing in elementary schools. Los Angeles, CA: Center for Research on Educational Standards and Student Tests, Graduate School of Education, UCLA.
- Stark, D. and Jennings, J. (2006). From the Capital to the Classroom. Year 4 of the No Child Left Behind Act. March, 2006. Retrieved June 2, 2007, from the EBSCOhost database.
- Sternberg, B. (2006). Real improvement for real students: test smarter, serve better. *Harvard Educational Review*, 76, 557-563
- U.S. Department of Education (2002). Elementary and Secondary Schools Act of 1965. Retrieved August 2, 2007, from U.S. Department of Education Web site: <http://www.ed.gov/legislation/FedRegister/proprule/2002-2/050602b.html>
- U.S. Department of Education (2004). *An Overview of NCLB*. Retrieved September 9, 2007, from U.S. Department of Education Web site: <http://www.ed.gov/print/nclb/overview/intro/execsumm.html>

Published by the Forum on Public Policy

Copyright © The Forum on Public Policy. All Rights Reserved. 2008.