

# AUDIT REPORT



CITY OF NEW YORK  
OFFICE OF THE COMPTROLLER  
BUREAU OF MANAGEMENT AUDIT  
**WILLIAM C. THOMPSON, JR., COMPTROLLER**

## **Audit Report on the Administration of New York State Standardized Tests by The New York City Department of Education**

*MD08-102A*

**July 22, 2009**



THE CITY OF NEW YORK  
OFFICE OF THE COMPTROLLER  
1 CENTRE STREET  
NEW YORK, N.Y. 10007-2341

---

WILLIAM C. THOMPSON, JR.  
COMPTROLLER

**To the Citizens of the City of New York**

Ladies and Gentlemen:

In accordance with the Comptroller's responsibilities contained in Chapter 5, §93, of the New York City Charter, my office has conducted an audit to determine whether the Department of Education (DOE) has adequate internal controls over the administering of New York State standardized tests for grades 3, 4, and 5.

DOE prepares students to meet grade level standards in reading, writing, and math and tests them to determine how well they are meeting mandated learning standards. Students in grades 3 through 8 take both the New York State standardized English Language Arts (ELA) Test and the New York State standardized Mathematics (Math) Test. Audits such as this provide a means of ensuring that City agencies adhere to relevant mandates and have adequate procedures and safeguards in place.

The results of our audit, which are presented in this report, have been discussed with DOE officials, and their comments have been considered in the preparation of this report.

I trust that this report contains information that is of interest to you. If you have any questions concerning this report, please e-mail my audit bureau at [audit@comptroller.nyc.gov](mailto:audit@comptroller.nyc.gov) or telephone my office at 212-669-3747.

Very truly yours,

A handwritten signature in cursive script that reads "William C. Thompson, Jr.".

William C. Thompson, Jr.

WCT/ec

**Report: MD08-102A**  
**Filed: July 22, 2009**

*Table of Contents*

**AUDIT REPORT IN BRIEF ..... 1**  
    Audit Findings and Conclusions..... 1  
    Audit Recommendations..... 2  
    Agency Response..... 2

**INTRODUCTION..... 3**  
    Background..... 3  
    Objective..... 5  
    Scope and Methodology ..... 5  
    Discussion of Audit Results..... 8

**FINDINGS AND RECOMMENDATIONS ..... 9**  
    Weaknesses in Monitoring System Controls over Test Administration..... 10  
        Ineffective Tracking System of Monitoring Visits and Monitoring Checklists ..... 10  
        Monitors Did Not Always Ensure that Test Administration Procedures Were Followed.... 11  
        Monitoring Checklists Were Inadequate ..... 12  
        Recommendations..... 14  
    DOE Does Not Perform Data Trend Analyses To Identify Possible Testing Irregularities ..... 15  
        Recommendations..... 17  
    Day Two and Day Three Portions of Tests Are at Risk of Manipulation..... 18  
        Recommendation ..... 18  
    DOE Does Not Have a Formalized Process to Ensure  
    Substantiated Allegations of Cheating Are Shared with OA..... 19  
        Recommendations..... 20  
    Other Issue ..... 21  
    Test-Administration Timeframe Set by the State Poses Security Risks for Exam Questions.. 21  
        Recommendation ..... 22

**APPENDIX**     Detailed Discussion of the DOE Response

**ADDENDUM**    DOE Response

*The City of New York  
Office of the Comptroller  
Bureau of Management Audit*

**Audit Report on the Administration of New York State  
Standardized Tests by the New York City  
Department of Education**

MD08-102A

---

**AUDIT REPORT IN BRIEF**

The New York City Department of Education (DOE) provides primary and secondary education to more than 1 million pre-kindergarten to grade 12 students in over 1,400 schools. DOE prepares students to meet grade level standards in reading, writing, and math and tests students to determine how well they are meeting these mandated learning standards.

Students in grades 3 through 8 take both the New York State standardized English Language Arts (ELA) Test and the New York State standardized Mathematics (Math) Test. This audit focuses on the administration of ELA and Math tests for students in elementary school grades 3, 4, and 5 only. The audit determined whether DOE has adequate internal controls over the administering of New York State standardized tests for grades 3, 4, and 5.

**Audit Findings and Conclusions**

DOE has adequate internal controls with respect to ensuring that schools are familiar with established procedures when administering the New York State standardized tests at elementary schools. In addition, the schools that we visited generally complied with the State testing guidelines, the DOE Handbook, and testing memoranda. However, DOE lacks sufficient preventive and detective controls aimed at deterring inappropriate manipulation of test scores, which would help to ensure the overall integrity of the assessment process.

DOE has established procedures for the administration of New York State standardized ELA and Math tests at elementary schools. DOE provides a Handbook and distributes test memoranda to its staff in an effort to keep them informed of all required procedures in administering State and Citywide tests. DOE also offers its staff annual training on proper methods in administering the tests as well as training of scoring staff to help identify testing irregularities when grading the long-answer portions of the exam.

We also found that for the most part, the schools that we visited complied with the State guidelines and the guidelines outlined in the Handbook. Our own review of the data and documentation collected by DOE for the 2007–2008 ELA and Math tests and our observations

conducted at the sampled schools on the day of testing did not reveal any instances of cheating. However, as more fully explained in the audit report, we cannot be assured that cheating did not occur.

Since achieving a positive school performance rating provides an added incentive for school officials to ensure that students perform well on standardized tests, there is a potential risk for inappropriate test manipulation. Based on our observations, we identified significant weaknesses that DOE has not addressed to help prevent or detect the manipulation of test scores. Specifically, DOE should improve its oversight of testing monitors to ensure that they are carrying out their duties properly and are using monitoring checklists more effectively. In addition, DOE should re-implement the use of analytics to identify possible testing irregularities and tampering and should institute stronger controls over the second and third sections of the tests. Finally, DOE should formalize a process to ensure that substantiated allegations of cheating are shared with the Office of Accountability (OA), the office primarily responsible for coordinating yearly testing and for compliance with New York State Education Department (NYSED) testing guidelines and DOE controls over the tests.

### Audit Recommendations

Based on our findings, we make 14 recommendations, 5 of which are listed below.

DOE should:

- Accurately track the assignment of testing monitors to ensure that they are being used effectively.
- Discuss with NYSED the possibility of obtaining the answer keys promptly after the administration of each test to enable DOE to perform a timely erasure analysis. However, DOE should perform erasure analysis to identify possible improprieties regardless of when it receives the answer key.
- Compile, maintain, and track data on the number of make-up exams that are taken for the Day Two and Day Three ELA and Math exams.
- Identify indicators to use in detecting unusual patterns that may be indicative of test tampering or irregularities and collect sufficient data to adequately track those indicators. Based on the information collected, DOE should target those schools with unusual patterns for further follow-up.
- Ensure that the Office of Special Investigations (OSI) formalizes a process to make certain that all instances of substantiated cheating are shared with OA, so that OA can strengthen existing controls or develop new ones in an effort to prevent cheating from occurring in the future.

### Agency Response

DOE officials generally agreed with the audit's recommendations but disagreed with one of them and did not address one of them. They also disagreed with the tone of the report. After carefully reviewing their comments, however, we found them to be without merit.

## INTRODUCTION

### Background

The New York City Department of Education (DOE) provides primary and secondary education to more than 1 million pre-kindergarten to grade 12 students in over 1,400 schools. DOE seeks to teach students to meet grade level standards in reading, writing, and math and then tests students to determine how well they are meeting these mandated learning standards.

Students in grades 3 through 8 take both the New York State standardized English Language Arts (ELA) Test and the New York State standardized Mathematics (Math) Test. This audit focuses on the administration of ELA and Math tests for students in elementary school grades 3, 4, and 5 only.

The ELA is a timed test that contains multiple-choice questions and performance assessment items. The test is given over a three-day period in grade 4 and a two-day period in grades 3 and 5. The Math test consists of two or three test sections, depending upon the grade level. Students answer multiple-choice, short response, and extended response questions. The first section of the test, given on Day One, is multiple-choice, and the second and third sections, given on Days Two and Three, consist of short answer and extended response questions.

After completion of the Day One ELA and Math tests, schools are required to deliver the student answer documents to the Integrated Service Center (ISC) in the respective school's borough by 3:00 p.m. The individuals delivering the documents are required to sign a log sheet to indicate their release of the documents to the ISC. Immediately after the answer documents are picked up from the ISC by a Scan Center courier, the log sheets are faxed to the Scan Center. The courier delivers the documents to the Scan Center in Long Island City, where the tests are logged in upon arrival and maintained until the multiple-choice answers are scanned. In those instances where school officials are unable to deliver the answer documents to the ISC by 3:00 p.m., they are required to deliver them directly to the Scan Center along with a letter from the school principal, explaining the reason for the delay.

The Day Two and Day Three portions of the ELA and Math tests are maintained at the schools until the completion of the make-up exams. The materials are required to be stored in a safe or storage vault and kept under strict security conditions. Afterwards, all the tests are picked up by the courier and delivered to a designated scoring site within each borough. Test materials are required to be secured at the scoring sites in locked rooms. Certified teachers, who are trained and required to pass a test prior to scoring, are selected to score the short and extended response questions at the scoring sites. Each exam is scored by more than one teacher, and teachers are not allowed to score exams from their own schools. The student answer results are then submitted to the State upon completion of the scoring process.

The student results of the ELA and Math tests are reported as scale scores and performance levels. With scale scores, the number of correct answers is converted to scores on a common scale so that achievement can be compared across grade levels. The student test results are also reported in four proficiency performance levels. The levels are as follows: 4—

meeting learning standards with distinction, 3—meeting learning standards, 2—partially meeting learning standards, and 1—not meeting learning standards.

Teachers and principals use the results of the ELA and Math tests to help them make decisions on whether a student will be promoted to the next grade and whether a student should receive academic intervention services.

The New York State Education Department (NYSED) sets general guidelines for testing and provides all school districts with the *School Administrator's Manual*. The manual provides guidelines to help ensure that tests are valid and equitable for all students. It includes information on testing accommodations, safeguarding test materials, administering make-up tests, and preparing test materials for scoring.

In an effort to standardize test administration procedures in all New York City classrooms, the Office of Accountability<sup>1</sup> (OA), the office primarily responsible for coordinating yearly testing and compliance with NYSED testing guidelines and DOE controls, developed a *Test Administration Handbook*. The handbook contains procedures for administering all State and Citywide tests and includes discussion of security issues and guidelines. Further, DOE develops and disseminates a test memorandum covering each test administration and information about test security processes and procedures. Each school also designates a Test Coordinator who is responsible for attending test administration meetings prior to each test administration and for distributing information regarding testing procedures.

Unannounced monitoring inspections before, during, and after the administration of State tests are conducted on a sample basis by DOE's Central Office staff and the State Education Department. The DOE monitors are required to fill out Test Administration Security Monitoring Forms (checklists), which are submitted to OA by the close of business on the day of each test.

If there are any suspected security violations during testing, they are required to be reported immediately to the City's Office of the Special Commissioner of Investigation (SCI). SCI has the discretion to investigate allegations of cheating or to forward the allegations to DOE's Office of Special Investigations (OSI) for further investigation.

Starting with the 2007–2008 school year, DOE gave school principals and their teaching teams broader discretion over, and accountability for, student achievement. The theory is that with the freedom to make decisions at the school level, school personnel are more accountable for their students' performance and progress. DOE has stated that schools that perform well will be rewarded, and schools that fail will face consequences. The results of State standardized test scores are important to DOE in its evaluation of student performance as well as in the assignment of school ratings. In addition, the results of the ELA and Math tests are important because they contribute to the assignment of certain accountability indicators to schools under both Federal and State Law, each of which has significant consequences, such as a requirement to plan instructional improvements, the mandatory provision to students of certain kinds of tutorial programs, and the potential for mandatory school restructuring or closure.

---

<sup>1</sup> At the end of the audit, DOE officials informed us that the Office of Accountability (OA) is now called the Division of Accountability and Achievement Resources (DAAR).

Achieving a positive correlation between a school's standardized test scores and its resulting performance rating provides an added incentive for school officials to ensure that their students perform well on standardized tests. It is, therefore, DOE's responsibility to implement adequate internal controls to ensure that its personnel adhere to State testing requirements to minimize the likelihood of inappropriate manipulation of the test-taking process.

### Objective

The objective of this audit was to determine whether DOE has adequate internal controls over the administering of New York State standardized tests for grades 3, 4, and 5 to ensure the integrity of the assessment process.

### Scope and Methodology

We conducted this performance audit in accordance with generally accepted government auditing standards. Those standards require that we plan and perform the audit in order to obtain sufficient, appropriate evidence to provide a reasonable basis for our findings and conclusions based on our audit objectives. We believe that the evidence obtained provides a reasonable basis for our findings and conclusions based on our audit objectives. This audit was conducted in accordance with the audit responsibilities of the City Comptroller as set forth in Chapter 5, §93, of the New York City Charter.

The scope period of this audit was the 2007–2008 school year.

To gain an understanding of the physical process used to administer the standardized tests, we interviewed officials from OA, including the Manager of Test Administration and Scanning, the Executive Director of Content and Assessment Support, the Director of Assessment, and the Assistant Manager of Test Administration and Scanning. We also met with the Director and Deputy Director of OSI.

To gain an understanding of the written policies and procedures used to administer the standardized tests, we reviewed the NYSED's *School Administrator's Manual* and DOE's *Test Administration Handbook* (Handbook) for grades 3 to 8, Test Administration Security Monitoring Form, and testing memoranda.

We conducted physical observations of the test-taking process at several judgmentally selected schools,<sup>2</sup> as shown in the following table. Our observations were conducted to test compliance with the controls that were described to us in the interviews or that were contained in policies and procedures.

---

<sup>2</sup> We targeted the sample to include schools from each of the City's four largest boroughs.



**Table I**

Schools Visited by Auditors during the  
Administration of Math and ELA Exams

<b>Test</b>	<b>School</b>	<b>Borough</b>	<b>Date of Observation</b>
ELA	PS230	Brooklyn	1/8/08
Math	PS112*	Queens	3/4/08
Math	PS130	Manhattan	3/5/08
Math	PS121	Bronx	3/5/08
Math	PS179	Brooklyn	3/5/08
Math	PS082	Queens	3/6/08
Math	PS097	Bronx	3/6/08
Math	PS234	Manhattan	3/6/08
ELA	PS111*	Queens	1/13/09
ELA	PS209	Brooklyn	1/13/09
ELA	PS001	Manhattan	1/14/09
ELA	PS048	Manhattan	1/14/09
ELA	PS253	Brooklyn	1/14/09
ELA	PS117	Queens	1/15/09
ELA	PS078	Bronx	1/15/09

\*These schools were selected because of their proximity to the Queens ISC and Scan Center.

In addition, we visited Public School 89 in the Bronx and Public School 15 in Manhattan during the administration of the Science test on April 30, 2008. Our review of the administration of the Science test included only these visits, which were conducted to observe compliance with the controls described to us in the interviews or contained in policies and procedures. We performed no other audit tests on Science exams since student performance measurements do not include proficiency in Science.

We visited the Queens ISC on March 4, 2008, and January 13, 2009, and observed the delivery of completed tests by Queens elementary schools and the pick-up of the tests by the courier service for delivery to the Scan Center. We also visited the Scan Center, observed the delivery of tests by the courier service, and conducted a walkthrough of the operations of the Scan Center. We also made two unannounced visits to the Manhattan scoring site during the scoring of the ELA and Math tests and met with the Assessment Implementation Director for Manhattan.

We obtained and reviewed copies of the checklists submitted by DOE monitors who visited elementary schools throughout the City during implementation of the 2007–2008 ELA and Math tests. The Director of Assessment Operations provided us with 147 checklists that were completed by monitors for the 2007–2008 ELA and Math tests. We determined whether the monitors reported that they arrived at the schools by 7:30 a.m., as required, and whether the checklists were completed.

We also reviewed the ISC log sheets for the 2007–2008 school year to identify elementary schools that were late in delivering the multiple choice portions of the ELA and Math tests to the ISC sites, along with required letters from the school principals. In addition, we reviewed information on the number of students who took ELA and Math make-up tests during the 2007–2008 school year to identify schools that had a large number of make-up tests.<sup>3</sup>

We obtained a listing of all substantiated testing cases involving allegations of cheating that were referred to OSI by SCI for school years 2005–2006, 2006–2007, and 2007–2008. We reviewed the OSI reports to determine which cases involved the State ELA and Math tests for elementary students in grades 3 through 5. We then determined the classification of the type of cheating that occurred and how it was identified.

We also attempted to assess possible indicators of the manipulation of tests through analytics. We reviewed the results of the State ELA and Math tests for grades 3 through 5 for the school years 2006–2007 and 2007–2008 for the schools in Districts 1 through 32. We calculated the difference between the mean scale scores for the two years. For both the ELA and Math tests, we then identified the three schools within each district that had the greatest increase in mean scale scores for each grade.

For these schools, we determined whether there were any substantiated allegations of cheating or a large number of make-up exams, which may have contributed to the schools' above average scores. We also determined for these schools whether the multiple-choice portion of the tests were delivered late to the Scan Center in the 2007–2008 school year. Since late delivery of an exam can provide school officials with time and opportunity to manipulate students' answer documents, it could contribute to the schools' above average scores. This test could be performed for only the 2007–2008 school year because, with the exception of the substantiated allegations, DOE maintains the documentation for make-up tests and late delivery of answer documents to the Scan Center for only one year. The documentation for the prior school years was destroyed.

For both the ELA and Math tests, we also identified the three schools within each district that had the greatest decrease in mean scale score for each grade. For these schools, we determined whether there were substantiated allegations of cheating in the 2006–2007 school year. We also determined whether a DOE monitor visited the schools during the administration of the tests that took place during the 2007–2008 school year. The presence of monitors could deter the occurrence of testing improprieties, which may have contributed to the schools' below average scores.

We attempted, on a limited basis, to see whether there was a correlation between an excessive number of erasures and student test scores. We determined the difference in the mean scale score between the 2006–2007 and 2007–2008 ELA tests for the fourth grade students for each school in Districts 1 through 32. We then selected the three schools with the highest increase in mean scale score for the fourth grade from each of the 32 districts. From this list of 96 schools, we randomly selected 15 schools using the Statistical Sampling System. We obtained

---

<sup>3</sup> We were able to review the checklists, ISC log sheets, and make-up test information for only the 2007–2008 school year because DOE destroys the records for prior school years.

the 2007–2008 ELA answer documents for one class of fourth grade students at each school to determine the number of erasures that were made in which the incorrect answer was changed to the correct answer.

### Discussion of Audit Results

The matters covered in this report were discussed with DOE officials during and at the conclusion of this audit. A preliminary draft report was sent to DOE officials and discussed at an exit conference held on May 4, 2009. On June 2, 2009, we submitted a draft report to DOE officials with a request for comments. We received a written response from DOE officials on June 16, 2009. In their response, DOE officials generally agreed with the audit’s recommendations but disagreed with one and did not respond to one. The response stated, “The Comptroller’s Draft Report identifies a number of potentially valid enhancements to existing DOE processes, which DOE has either already adopted or is considering.”

However, the DOE response, which is lengthy, included objections to our findings. After carefully reviewing DOE’s arguments, we found them to be without merit. DOE’s comments appear to be an attempt to divert attention from the fact that DOE’s internal controls over standardized tests can be improved. By implementing or considering implementing 12 of the 14 recommendations, DOE officials confirm the benefit of this audit and their desire to improve the internal controls over standardized tests.

A detailed discussion of the DOE response is included as an appendix to this report, and the full text of the DOE response follows the appendix as an addendum.

## FINDINGS AND RECOMMENDATIONS

DOE has adequate internal controls with respect to ensuring that schools are familiar with established procedures when administering the New York State standardized tests at elementary schools. In addition, the schools that we visited generally complied with the NYSED testing guidelines and the DOE Handbook and testing memoranda. However, DOE lacks sufficient preventive and detective controls aimed at deterring inappropriate manipulation of test scores, which would help to ensure the overall integrity of the assessment process.

DOE has established procedures for the administration of New York State standardized ELA and Math tests at elementary schools. DOE provides a Handbook and distributes test memoranda to its staff in an effort to keep them informed of all required procedures in administering State and Citywide tests. DOE also offers its staff annual training on proper methods in administering the tests, as well as training of scoring staff to help identify testing irregularities when grading the long answer portions of the exam.

We also found that for the most part, the schools that we visited complied with the State guidelines and the guidelines outlined in the Handbook. Those guidelines include ensuring that test booklets are stored in secure locations, that there are hall proctors, that there are separate rooms for students arriving late on the day of the exam and for students not being tested, that classrooms are adequately set up with all visual aids covered, and that teachers sign the test security material forms before and after the tests.<sup>4</sup> Our own review of the data and documentation collected by DOE for the 2007–2008 ELA and Math tests and our observations conducted at the sampled schools on the day of testing did not reveal any instances of cheating. However, as more fully explained in the audit report, we cannot be assured that cheating did not, in fact, occur.

Since achieving a positive school performance rating provides an added incentive for school officials to ensure that students perform well on standardized tests, there is a potential risk for inappropriate test manipulation. Based on our observations, we identified significant weaknesses that DOE has not addressed to help prevent or detect the manipulation of test scores. Specifically, DOE needs to improve its oversight of testing monitors to ensure that they are carrying out their duties properly, and ensure that monitoring checklists are used more effectively. In addition, DOE should re-implement the use of analytics to identify possible testing irregularities and tampering and should institute stronger controls over the second and third sections of the tests. Finally, DOE should ensure that substantiated allegations of cheating are shared with OA, the office primarily responsible for coordinating yearly testing and compliance with NYSED testing guidelines and DOE controls over the tests. These issues are discussed in more detail in the following sections of the report.

---

<sup>4</sup> Test security material forms are required to be signed before and after the administration of the test by each teacher to verify that the correct number of exams were delivered and returned.

## Weaknesses in Monitoring System Controls over Test Administration

The monitoring system that DOE has in place for the administration of the standardized tests is inadequate. In addition to its monitors, DOE identified the Test Administration Security Monitoring Forms (checklists) used by the monitors as one of its internal controls over the administration of the State tests. DOE officials stated that unannounced monitoring inspections are conducted before, during, and after testing to deter cheating. However, DOE does not track monitoring visits or the submission of monitoring checklists. In addition, the monitors did not always ensure that schools adhered to test administration procedures. Furthermore, the content of the monitoring checklists needs improvement.

### Ineffective Tracking System of Monitoring Visits and Monitoring Checklists

DOE does not keep track of the monitors assigned to visit schools and the submission of checklists. DOE provided us with a schedule of monitors and the schools they were assigned to visit during the administration of the 2007–2008 ELA and Math exams. According to the schedule of monitors, there were 265 monitoring visits scheduled for those tests. We requested all the monitoring checklists for schools visited by monitors for the 2007–2008 ELA and Math tests and were provided with only 147 checklists. Moreover, only 128 of these checklists matched the schools listed on the monitoring schedule. We determined that nine of the monitors visited schools that were different from those listed on the schedule, including seven of the eight monitors who accompanied us on our site visits to schools where testing was taking place. In addition, 10 checklists were submitted for visits that were not listed on the monitoring schedule. Therefore, we did not receive monitoring checklists for 128 (48%) scheduled visits.

As previously stated, based on DOE’s schedule of school visits, DOE intended to have 265 monitoring visits conducted. However, it did not track the actual monitoring visits or the submission of monitoring checklists. Nor did DOE maintain an official record of the actual monitoring visits. As a result, DOE was unable to determine whether monitors were present at the schools for which there were no checklists or whether the monitors visited the schools but their checklists were missing. After being told of our findings, DOE officials provided us with a list that identified monitoring visits for which checklists were submitted and monitoring visits that were confirmed but for which no checklists were available. DOE officials informed us that they had an employee contact “all of the individuals on the list of prospective monitors for whom monitoring forms were not available to confirm which ones had, and had not been able to serve as monitors in 2008.” Based on these contacts DOE officials claim that an additional 34 monitoring visits were conducted; however, DOE was unable to provide the monitoring checklists for these visits.

When we informed DOE officials about the lack of checklists, they stated that the monitors are full-time DOE employees with other duties and that although they are scheduled to assist with monitoring, they sometimes need to cancel because of other job responsibilities. In addition, a DOE official stated that the checklists were reviewed at the time the checklists were actually submitted by the monitors, but that no record-retention process was in place during the

2007-2008 school year. The official also stated that DOE currently will now implement a record-retention process for the scheduling of monitors and the submission of checklists.

DOE officials also stated that the checklists are not the primary mechanism for reporting violations identified during monitoring. DOE claimed that monitors are explicitly required to notify OA immediately if they identify any potential testing irregularities during visits. We requested from DOE all instances during the 2008 school year in which monitors reported potential testing irregularities for grades 3, 4, and 5. DOE provided us with evidence of only one instance. We are not certain whether additional potential testing irregularities were identified by monitors, because OA does not keep a log of irregularities reported by monitors.

#### Monitors Did Not Always Ensure that Test Administration Procedures Were Followed

Monitors did not always arrive at the schools at the mandatory time to observe all required pre-administration procedures. As a result, they could not ensure, among other things, that the test booklets were adequately safeguarded prior to the test.

DOE's checklist states, "Monitors must arrive at assigned school by 7:30 a.m." It is important for monitors to arrive by 7:30 a.m. to ensure that the test booklets are secured in a locked place and are still in the shrink-wrapped packaging. To maintain test security and to prevent advance review of the test questions, DOE requires that the shrink-wrapped packages not be opened until 60 minutes before administering the test. We reviewed the 147 monitoring checklists that we received for the 2007–2008 ELA and Math tests and determined that 47 (32%) monitors arrived at the schools after the required time of 7:30 a.m. Moreover, 24 (16%) of the 147 monitors arrived at or after 8:00 a.m., which depending on the start time of the test could be a problem for the monitor to determine whether tests had remained in the shrink-wrapped packaging until one hour prior to the test start time, which is no later than 9:15 a.m. In fact, one monitor arrived as late as 10:20 a.m. Furthermore, the checklist completed by this monitor contained a note indicating that the monitor was not present at the school for the start of the exam and was, therefore, unable to determine whether the test booklets were still in the shrink-wrap packaging on the day of the test. However, this monitor marked "yes" to the question on the pre-administration section of the checklist that asked whether tests booklets were stored in a secure place. Since the monitor arrived at 10:20 a.m., it would have been impossible for the monitor to observe the pre-administration storage of the exams.

As stated earlier, there were eight monitors who accompanied us on our observations of the 2007–2008 ELA and Math test. Only six of these eight monitors submitted the required checklists. Three of the eight monitors did not arrive at the schools by 7:30 a.m. as required by DOE. In addition, two of the eight monitors did not observe the testing classrooms at all to ensure that the exam was administered according to State guidelines. Further, three of the eight monitors observed the testing classrooms for only a portion of the exam, not the entire duration of the exam.

During our own observations of the administration of the tests, we found it necessary to observe the testing classrooms for the duration of the exam. Our observations were to determine whether the teachers recorded start times on the chalk boards, whether materials that might assist

students on the test were covered, whether students were seated so that they could not see answer sheets of others, and whether students were given the appropriate amount of time for the test. DOE's checklists ask specific questions that require the monitors to observe the classrooms while the test is in progress. Nevertheless, during our observations we noted that some monitors did not observe the classrooms while the test was in progress. Consequently, we question how the monitors were able to fully and accurately complete the checklists without observing the classrooms during testing.

In February 2009, as we were completing the audit fieldwork, we informed DOE officials of these issues concerning the monitors. The following day, DOE officials stated that the monitors "were not assigned to those schools on those days to perform a monitoring function, and thus would not have been expected to complete a monitoring checklist." During the course of the audit, DOE officials did not make this assertion, even when we made our request for the monitoring checklists completed by these individuals. It is doubtful that DOE would accept the monitoring checklists and present them to us as such if the individuals who completed them were not sent to the schools to perform a monitoring function. In addition, we observed that several individuals clearly acted in the capacity of monitors. For example, they asked questions of the school administrators, observed testing classrooms, and filled in sections of the monitoring checklist. Moreover, one individual informed a teacher that certain information in the classroom needed to be covered during the test.

Seven of the eight monitors accompanying us were scheduled to visit different schools, which DOE also used as the basis for its argument that the monitors should not be considered actual monitors but rather individuals assisting us during our site visits. At the exit conference, DOE officials stated that these individuals were sent to "assure that the auditors themselves observed NYSED test administration restrictions." However, we doubt this assertion since the majority of the monitors did not even accompany us as we patrolled the hallways to assess compliance with test administration policies.

The presence of monitors is an effective tool that can be used to minimize the potential for improper conduct. Since monitors are an important part of DOE's efforts to identify testing irregularities, it is vital that DOE track monitoring activities and review the checklists submitted to identify testing irregularities. For example, one allegation of cheating substantiated by OSI found that a teacher allowed students additional time to complete the test and failed to start one of the tests on time. The presence of monitors who are actively observing classrooms and effectively monitoring tests is meant to prevent these types of improprieties from occurring.

#### Monitoring Checklists Were Inadequate

The checklist used by the monitors was inadequate and needed to be modified to reflect only what can be ascertained from actual observations. Some of the items on the checklist could not be answered by the monitor based on observations, thereby rendering the answers of little or no value. Consequently, the checklists used by the monitors were not as effective as they should have been in identifying weaknesses in the administration of the tests.

The purpose of the monitors' presence during the administration of the tests is to determine whether the schools are adhering to State guidelines. They are able to do so through patrolling the hallways and observing the testing classrooms. The monitors are not allowed to enter the classrooms or cause any distractions while the exam is taking place. The checklist is a tool that should be used by DOE to identify weaknesses in school administration of the test and noncompliance with State guidelines. However, the checklist should be modified to include only those questions that can actually be answered by the monitors, and the monitors need to complete the checklists accurately. We found that more than 21 percent (6 of the 28) of the questions on the checklist required answers by the monitor that would be impossible to determine through routine observations. In fact, the monitors would have had to be present in the classrooms to answer five of the six questions. The six questions are as follows:

- Students have been reminded to bring #2 pencils to class.
- The teacher follows the detailed directions in the *Teacher's Directions*.
- The teacher walks around the room ensuring that all students appear to understand the directions and are properly marking the answer sheets.
- Test taking procedures are explained again to any student who appears to be having difficulty.
- Students erase all extraneous pencil marks and darken bubbles on answer documents when necessary.
- Teachers DO NOT give help on specific questions, translate, or tell a student to review an answer.

Although these questions cannot be answered by the monitors through observations alone, 104 (71%) of the 147 monitoring forms we reviewed had a "yes" response for all questions, including these six. (Generally, the remaining 43 forms had one or more responses of "Not Applicable.") In addition, the checklist required monitors to ensure that shrink-wrapped packages of test booklets were not opened until the day of the test. The checklist did not indicate that the shrink-wrap cannot be opened until 60 minutes prior to the test administration, although this is required by DOE procedures which state that "testing coordinators may not open shrink-wrapped packages . . . until 60 minutes before test administration." It is unclear from the checklist whether the monitors were ensuring adherence to this requirement, since the checklist did not specify how long before the test administration the shrink-wrap could be opened. Therefore, a monitor could record a "yes" response to this question even if the shrink-wrap was opened hours before the test administration, which could compromise the integrity of the exam. Consequently, the checklists may have provided DOE with incorrect information and may have inaccurately indicated that schools were in compliance with State test-taking guidelines.

Once informed of our findings, DOE officials revised the checklist. In addition to modifying some questions and including additional ones, the six questions about which we expressed concern were removed from the checklist.



## Recommendations

DOE should:

1. Accurately track the assignment of testing monitors to ensure that they are being used effectively.
2. Ensure that all monitors submit a monitoring checklist. DOE should consider requiring the monitors to submit a copy of the checklist to OA electronically so that the information can be maintained for more than a year.

**DOE Response:** “DOE has either already adopted or is considering . . . enhancing monitoring checklist retention and scheduling processes (implemented winter 2009) [and] storing returned monitoring checklists electronically (under consideration).”

3. Ensure that monitors arrive at the schools by 7:30 a.m. on the day of the tests to make certain test booklets are properly safeguarded before the administration of the exam.

**Auditor Comment:** DOE did not address this recommendation in its response.

4. Ensure that monitors are trained to understand the importance of their functions and actively observe testing classrooms to ensure compliance with guidelines.
5. Ensure that the revised checklist is used by the monitors and should modify the checklist to indicate clearly that the shrink-wrapped test booklets should not be opened until one hour prior to the test administration.

**DOE Response:** “DOE has . . . already adopted . . . clarifying monitoring checklist items and clarifying monitor tasks [and] adding a statement to the monitoring checklist instructions that mirrors statements in materials provided to schools and monitors that test-booklet shrink wrap may not be opened more than one hour before the exam (implemented winter 2009).”

**Auditor Comment:** We are pleased that DOE has agreed to implement these recommendations; however, the revised monitoring checklist that DOE provided as Appendix C of its response states, “Building Supervisors/Testing Coordinators may open shrink-wrap packages 60 minutes before test administration.” We believe that the instructions should be clearer and in line with DOE’s own Assessment Memorandum, which states, “School Supervisors/Testing Coordinators may not open shrink-wrapped packages in order to complete class sets until 60 minutes before test administration.” (Emphasis in original.)

## DOE Does Not Perform Data Trend Analyses To Identify Possible Testing Irregularities

DOE stopped using computerized data trend analyses to identify patterns and irregularities of test scores among schools. As a result, the agency is unable to identify trends that may indicate inappropriate manipulation of test scores or cheating.

Computerized data trend analyses, although not required by NYSED, would clearly assist DOE in identifying test program strengths and specific testing program vulnerabilities. This information would allow DOE to strengthen internal controls in the areas where weaknesses are identified and help DOE in targeting the schools that should be monitored, rather than relying on DOE's current practice of randomly selecting schools.

According to DOE officials, prior to calendar year 2002, DOE performed erasure analyses of answer documents to identify patterns and irregularities that may indicate cheating. Erasure analysis looks for an abnormal number or pattern of changed answers (from incorrect to correct) on student answer documents. Although excessive erasures are not necessarily indicative of cheating, erasures can be a sign that further investigation is required. The identification and investigation of the erasures can determine whether patterns exist among students, classes, grades, or schools.

Our own erasure analysis of the test answer documents from one fourth-grade class from each of the 15 sampled schools highlighted two students from two different schools with an excessive pattern of erasures.<sup>5</sup> One student changed the incorrect answers to the correct answers for 12 of 15 erased responses (out of 28 exam questions). Another student changed the incorrect answers to the correct answers for all 16 of the erased responses. Although these two students were not in the same class or school, the excessive number of erasures from incorrect to correct responses warrants further investigation. While our analysis was limited, if DOE were performing erasure analysis, it would be able to do so for all exams using erasure analysis software.

DOE officials stated that erasure analysis is used when appropriate during the course of specific investigations of testing irregularities. (Since DOE provided no supporting evidence of this, we are unable to ascertain the extent to which erasure analysis is currently being used.) However, officials stated that this analysis is not performed for all ELA and Math exams. When we first questioned DOE as to the reason it is not performed more widely, officials first contended that the State stopped sharing the answer key at the time of the exam and the answer key is required to perform erasure analysis. However, the State stopped sharing the answer key in 2006 and, according to DOE, erasure analysis has not been performed since the tests that were scored in 2001. Officials later stated that they spoke with an individual who was in charge of the Division of Assessment and Accountability between Fall of 2001 until she left the DOE in 2006 who claimed, according to DOE, that patterns revealed by erasure analyses of tests "were not resulting in substantiated cases of testing misconduct." However, we could not verify this claim because neither

---

<sup>5</sup> In conducting our erasure analysis, we manually reviewed the 336 student ELA answer documents pertaining to our sample of 15 schools and noted 135 answer documents in which students had erased wrong answers and changed them to the correct ones. Out of this population, we noted the two students who had an excessive number of erasures.

DOE nor this former official provided any evidence (e.g., results of previous erasure analysis performed and subsequent investigations) to substantiate this assertion.

Even if this claim were true, however, the incentive to manipulate test scores, and the risk that this may occur, has increased in recent years. Starting in 2006, under the mandate of the federal No Child Left Behind Act, New York State rather than the City has oversight of these tests. In addition, starting with the 2007-2008 school year, DOE gave principals and their teaching teams broader discretion over, and accountability for student achievement. In such an environment, tools (such as erasure analysis) should be implemented to help DOE more effectively identify possible testing irregularities.

In addition to erasure analyses, there are other indicators that may signal testing irregularities. These include, but are not limited to, excessive number of make-up exams, late delivery of exams to the ISC and Scan Center, and a significant increase in test scores. Students absent on the day of the exam take a make-up exam that is identical to the exam administered earlier.<sup>6</sup> Students who have already taken the exam and school officials have, therefore, already had access to the exam questions, which provides the opportunity to give an unfair advantage to students taking make-up exams. This could create a situation in which poorly performing students are asked to stay home on the day of the exam so that they can be assisted later in the taking of the make-up exam. Consistent late delivery of tests could also be an indication of test tampering, since late delivery of an exam can provide school officials with the time and opportunity to review and correct students' answer documents. Significant increases in test scores may also be indicative of test tampering, especially if accompanied by other indicators (e.g., high number of make-up exams).

We attempted, but were unable, to track the late delivery of answer documents and the number of make-up exams taken by schools over the last few years because DOE only maintains this information for approximately one year from the date of the exam. Therefore, we were unable to draw any conclusions from our attempted data-trend analysis due to the very limited information available. With respect to make-up exams, DOE maintains data on the number of make-up exams taken for only Day One of the ELA and Math exams, and likewise maintains this information for just one year. DOE does not track the number of Day Two or Day Three make-up tests for either the ELA or the Math exams. Without this information, DOE does not have a true number of make-up exams that are actually taken. Without the tracking of late delivery of answer documents, the number of make-up exams, significant increases in test scores, and other useful mechanisms that may highlight potential testing irregularities, DOE can not ensure that cheating does not exist.

DOE officials acknowledged that the agency “does not conduct data trend analyses for purposes of identifying testing irregularities.” They further stated, “The DOE does not currently maintain longitudinal data nor conduct data trend analyses on make-up exams [or] late delivery of exams. . . . With the 2009 administration we intend to begin tracking these data longitudinally so that we can begin these analyses.”

---

<sup>6</sup> It is NYSED policy to provide only one form of the exam.

## Recommendations

DOE should:

6. Discuss with NYSED the possibility of obtaining the answer keys promptly after the administration of each test to enable DOE to perform a timely erasure analysis. However, DOE should perform erasure analysis to identify possible improprieties regardless of when it receives the answer key.

**DOE Response:** “DOE . . . is considering . . . conducting cost-effective methods of erasure analysis.”

**Auditor Comment:** DOE acknowledges that this recommendation is a valid enhancement. We, therefore, urge DOE to implement this recommendation without delay.

7. Compile, maintain, and track data on the number of make-up exams that are taken for the Day Two and Day Three ELA and Math exams.
8. Identify indicators to use in detecting unusual patterns that may be indicative of test tampering or irregularities and collect sufficient data to adequately track those indicators. Based on the information collected, DOE should target those schools with unusual patterns for further follow-up.

**DOE Response:** DOE is considering implementing recommendations 7 and 8 and stated, “DOE . . . is considering . . . analyzing longitudinal data related to late deliveries and make-up exams.”

**Auditor Comment:** In its response, DOE acknowledges that these recommendations are valid enhancements. We, therefore, urge DOE to implement these recommendations without delay.

9. Maintain and track all indicators necessary for preventing or detecting testing irregularities for multiple years.

**DOE Response:** DOE partially agreed to this recommendation stating, “DOE . . . is considering . . . analyzing longitudinal data related to late deliveries and make-up exams.”

**Auditor Comment:** In addition to maintaining data related to late deliveries and make-up exams, DOE should also maintain and track information on other useful mechanisms that may highlight potential testing irregularities.

Day Two and Day Three Portions of Tests  
Are at Risk of Manipulation

The Day Two and Day Three portions of the ELA and Math tests remain at schools for a number of days, thereby, increasing the risk of their being inappropriately manipulated.

After completion of the Day One sections of the ELA and Math tests, schools are required to deliver the student-answer documents to the ISC in the school's borough by 3:00 p.m. However, the Day Two and Day Three sections of the ELA and Math tests are maintained at the schools until the completion of the make-up exams—up to six or seven days from the initial administration of the exam—at which time they are picked up by a courier.

DOE requires the immediate delivery of the Day One section of the tests because the multiple choice answers are easily susceptible to inappropriate manipulation. However, given the time the completed Day Two and Day Three sections of the exam remain at the schools, they are also susceptible to inappropriate manipulation. Although DOE requires that schools store the completed tests in a securely locked place until they are picked up by the courier, they are still in the schools' possession. Furthermore, DOE does not currently have a mechanism in place to ensure that schools comply with the requirement that tests are kept secure and locked. Since the tests are maintained at the schools for a number of days, there is an opportunity for answers to be reviewed by school staff or administrators and changed from incorrect to correct answers. This possibility was confirmed to us by the Assessment Implementation Directors during our visits to the scoring sites. They stated that they have identified instances of adult handwriting appearing on the Day Two and Day Three test booklets. In addition, the potential exists that teachers or administrators can review the tests and return them to students who did poorly so that answers can be revised. In fact, according to an OSI investigation, a monitor was sent to a school and witnessed a teacher in her office preparing to review the previous day's math exam with three students. The boxes that contained the math tests were not properly safeguarded in a locked place.

To minimize these risks, DOE must improve the internal controls over the storage of completed Day Two and Day Three sections of the tests. For example, DOE could consider a procedure whereby all Day Two and Day Three tests and answer documents are stored in boxes with tamper proof, dated seals, and it should randomly assign monitors to visit schools to see that the boxes are appropriately sealed.

Recommendation

10. DOE needs to institute additional controls to decrease the risk of inappropriate manipulation of answer documents for the Day Two and Day Three portions of the standardized tests.

**DOE Response:** "DOE . . . is considering . . . using 'dated seals' on Day Two and Day Three materials."

**Auditor Comment:** In its response, DOE acknowledges that this recommendation is a valid enhancement. We, therefore, urge DOE to implement this recommendation without delay.

## DOE Does Not Have a Formalized Process to Ensure Substantiated Allegations of Cheating Are Shared with OA

DOE does not have a formalized process in place to make certain that substantiated cases of cheating on State tests, whether investigated by SCI or DOE's OSI, are shared with OA, the office primarily responsible for ensuring that test administration is in compliance with NYSED and DOE guidelines. Consequently, we cannot be assured that OA is aware of problematic areas that need to be addressed when administering the standardized tests.

All allegations of cheating on DOE tests are reported to SCI. SCI can choose to investigate an allegation or refer it to DOE's OSI for investigation. If SCI substantiates an allegation, it shares its report with OSI so that the recommendation for disciplinary action can be carried out. In instances where cases are referred to OSI, OSI is the unit responsible for investigating the complaint and determining whether it is founded. Although OSI is aware of all substantiated allegations of cheating, this information is not shared with OA.

For example, our review of the 13 allegations substantiated by OSI during 2006 through 2008<sup>7</sup> revealed 9 (69%) instances of improper proctoring in the classroom by teachers and school officials. Improper proctoring may include proctors gesturing that a response is correct or incorrect, writing on the blackboard, orally providing the answers, giving the tests back to specific students to re-check, changing a response on behalf of the students, and providing definitions. Our review of the substantiated allegations of improper proctoring also involved a teacher instructing students to pay attention to a particular question upon the teacher finding that a number of students had answered the question incorrectly. In addition, there were instances where the tests were easily accessible because they were left in an unsecured location.

OSI should formalize a process by which all substantiated instances of cheating are shared with OA since OA is responsible for ensuring compliance with NYSED testing guidelines and DOE controls. This information would allow OA to develop ways to strengthen its controls to prevent cheating from occurring. For example, since improper proctoring was identified as a significant issue in 9 out of 13 substantiated instances of cheating, OA needs to develop controls to address it, such as assigning proctors, other than the class's own teacher, to monitor the exams. Had OSI shared the substantiated cases of cheating with OA, OA may have strengthened existing controls or designed additional ones. Moreover, as a deterrent, DOE should remind teachers of the disciplinary consequences they can face if allegations against them are found to be substantiated.

At the exit conference, DOE officials stated that OA is aware of allegations during the actual conduct of the SCI and OSI investigations because once either SCI or OSI is notified of an allegation that office works closely with an OA specialist in the process of investigating the allegations. After the exit conference, DOE officials provided us with a written document that states that when the investigator "learns of a case from OSI, she feeds back the information to her supervisor in [OA] and proposes changes to test security processes whenever it is determined that revised procedures might reasonably improve the reliability of test administration, handling or

---

<sup>7</sup>Our review of OSI's investigation of testing irregularities did not include allegations that were unsubstantiated. In addition, our review covered those irregularities that were referred to OSI for Math and ELA exams for grades 3, 4, and 5 only.

scoring.” Nevertheless, DOE officials acknowledged that a process for sharing allegations with OA was not formalized. Officials then stated that effective March 2009, all substantiated OSI reports will be shared with OA.

After the exit conference DOE officials provided us with 15 information sheets that indicated collaboration between OSI and the OA specialist regarding investigations of alleged testing irregularities reported in 2008. We requested the steps or changes OA has taken with regards to testing procedures as a result of the specialist’s involvement in these investigations. After the exit conference, the OA specialist informed us that changes were made to the Test Material Security form as a result of an allegation, but she was unable to recall when the change took effect.

### Recommendations

DOE should:

11. Ensure that OSI formalizes a process to make certain that all instances of substantiated cheating are shared with OA so that OA can strengthen existing controls or develop new ones in an effort to prevent cheating from occurring in the future.

**DOE Response:** “DOE has . . . already adopted . . . formalizing process by which DAAR [Division of Accountability and Achievement Resources] obtains final reports of substantiated testing allegations from OSI (OSI reports) and ATU (SCI reports).”

12. Require that schools assign exam proctors other than the class’s own teacher, so that teachers do not proctor the exams for their own classes.

**DOE Response:** “The Comptroller’s proposal to require 7- and 8-year old children to take state high-stakes tests in rooms where no familiar adult is present is not in the best interest of children and has no basis in the governing NYSED rules.”

**Auditor Comment:** Sixty-nine percent of the allegations substantiated by OSI from 2006 through 2008 involved improper proctoring in the classroom by teachers and school officials, confirming that this is one of DOE’s biggest concerns. Assigning exam proctors other than the classrooms’ own teacher minimizes the risk that improper proctoring may occur. We seriously doubt DOE’s assertion that students almost half way through the school year would not be familiar with the teachers from across the hall or the room next door.

13. Periodically inform teachers and administrators of the possible consequences they can face if they are found to be involved with test tampering.

**DOE Response:** “Teachers and administrators are made fully aware of proper proctoring guidelines as a result of DOE’s supplemental control requiring each proctor to review the Test Administration Handbook and sign a statement confirming their understanding. Teachers and administrators know that if they are found to have engaged in proctoring irregularities, they face serious disciplinary action, including termination.”

***Auditor Comment:*** Our recommendation is that DOE inform teachers of the consequences they can face if found to be involved with test tampering, even though the Test Administration Handbook does not specify consequences. DOE should prepare a formal written document that identifies the possible consequences and ensure that it is distributed to all teachers and administrators.

## Other Issue

### Test-Administration Timeframe Set by the State Poses Security Risks for Exam Questions

The State does not require all counties within New York to administer tests on the same dates, but rather establishes a timeframe during which tests are to take place. NYSED defines the standardized assessment timeframes. DOE cannot deviate from this timeframe. However, the flexibility allowed to school districts throughout the State to decide when to administer the tests weakens the effort to maintain secure tests. The City has the discretion of deciding which days to administer the exam within the period established by the State. There is a potential for the security of the exam to be compromised, since the exam is identical throughout the State and it is given at various times in different parts of the State. Individuals with earliest access to the exam have an opportunity to provide students or others elsewhere with information about the exam.

The timeframe allowed for administration of the tests was brought up as a concern by officials at schools that we visited. Three assistant principals and one Testing Coordinator at four different schools that we visited stated that the window of time allowed for the exams jeopardizes the security of the exams. In addition, in an allegation of cheating investigated by OSI, a student from a different part of the State admitted to her teacher that she had received access to the third grade ELA exam from her mother, a teacher in the New York City Department of Education. The reverse of this situation can also occur, whereby, City students or schools can receive access to tests that have already been administered in other areas of the State. Although the time period for administering the exam is established by the State, under the current testing guidelines there is a possibility that information about the tests can be shared. We discussed this issue with DOE officials, and they acknowledged it as a concern.

In a response provided at the exit conference, DOE officials stated, “DOE has made it a local requirement that DOE schools administer at the very beginning of the test window.” However, this practice does not eliminate the possibility of City teachers` sharing information with teachers or students from districts outside the City. They also stated, “DOE has long been on record with NYSED that the state should shorten the test window.” However, we requested, but never received, any evidence to support this statement.



Recommendation

14. DOE should work with the State to address and minimize the opportunity for sharing of test information throughout the State that is posed by the test-administration timeframe.

***DOE Response:*** “DOE has . . . already adopted . . . asking NYSED to shorten the test administration window.”

***Auditor Comment:*** DOE has stated that this recommendation was implemented in the last several years but when we requested evidence to support this claim DOE did not provide it.

## DETAILED DISCUSSION OF THE DOE RESPONSE

During the course of the audit, we had numerous meetings and correspondence with DOE officials to discuss the issues addressed in this report. Nevertheless, in its response, DOE strongly objected to our methodology and our findings. We disagree with DOE officials' arguments and, therefore, have added this Appendix to record the main issues raised in the DOE response and our responding comments. (For the full text of DOE's response, see the Addendum to this report.)

### **Re: Overall Objections to the Report**

#### **DOE Response**

Eighteen months after the audit began, the Comptroller is finally releasing his report, which ignores his auditors' entirely positive findings about the integrity of the tests and their results and makes a series of trivial debater's points, replete with clear mistakes and loud demands for changes DOE already made months ago.

#### **Auditor Comment**

From its response, it is apparent that DOE does not understand the significance of the audit's findings. Although the schools visited generally complied with the NYSED testing guidelines, the audit found that DOE lacks sufficient preventive and detective controls aimed at deterring inappropriate manipulation of test scores. Apparently DOE does not think these controls are important, calling them "trivial debater's points." We disagree. As we state in the report, starting with the 2007-08 school year, DOE gave schools increased accountability over student achievement. One of the indicators used to assess this is student performance on standardized tests. DOE attempts to downplay that increasing the incentive of schools to ensure that their students perform well likewise increases the risk of inappropriate test manipulation. This is a concern to us. Controls to prevent or detect irregularities are vital in ensuring the overall integrity of the assessment process. Accordingly, DOE should reconsider its position regarding the audit's overall findings.

Regarding changes that DOE claims were made months ago, the agency provided no evidence that any of these changes were implemented prior to our discussing the audit's findings with DOE personnel.

#### **DOE Response**

Instead of confirming that his auditors found no problems on more than a million student exams and proposing modest improvements, the report complains about the way that DOE *used to* implement a *single* procedural control that DOE has *voluntarily* adopted—above and beyond anything federal and state law require and any other district in the country uses—which the Comptroller *never actually observed in action* and *badly*

*misrepresents*. And instead of celebrating test score gains that the audit has exhaustively validated across two years and hundreds of thousands of children, the report levels last-minute, irresponsible accusations against two fourth graders for making—and correcting—the most common of ‘bubbling’ mistakes (each obviously got off by a row or two in moving answers from test booklet to bubble sheet). [Emphasis in original.]

### **Auditor Comment**

Our review of DOE’s use of testing monitors is an accurate representation. DOE provided no evidence that it had made any changes in how monitoring was conducted prior to our discussion of the audit findings with agency personnel. In addition, DOE’s statement that we have not observed test monitoring in action is categorically incorrect. The individuals who were present at the schools on the days of our observations were presented to us as monitors. Regarding DOE’s other points, the audit did not include the review of more than a million student exams or the validation of test score gains. Finally, the report made no accusations against the two students, but rather suggested that further investigation was necessary.

### **Re: Audit Objective**

#### **DOE Response**

In December 2007, the Comptroller notified DOE of his intent to audit DOE’s administration of state English and Math tests during the 2007-08 school year to consider the integrity of the examination process and look for evidence of “cheating.” Although the Comptroller’s formal audit notification was limited to test administration during the winter 2008 English and Math tests, the Comptroller subsequently and without notification expanded the scope of the audit to include the April 2008 state Science tests, and expanded it again to include the 2009 state English tests.

#### **Auditor Comment**

DOE misstates our audit objective numerous times in its response. The objective of the audit was not to look for evidence of cheating, but was and has always been to determine whether DOE has adequate internal controls over the administration of New York State standardized tests to ensure the integrity of the assessment process. Further, the formal audit notification that DOE refers to did not mention or limit the time period or the specific tests to be reviewed.

### **Re: Evaluation of DOE Controls over Administration of Tests**

#### **DOE Response**

DOE has adopted over 20 supplemental controls to assure the integrity of test results. DOE undertakes each control voluntarily, above and beyond state requirements, and supplementary to ongoing oversight both by NYSED and by

two investigative agencies (the Office of the Special Commissioner of Investigation [SCI] and the Office of Special Investigations [OSI]). The single supplemental control the Comptroller looked at is DOE's unannounced monitoring of a sample of schools . . . which DOE voluntarily conducts to augment New York State's own unannounced monitoring.

As additional precautions, DOE voluntarily implements a set of procedures, not required by federal or state law, to verify schools' adherence to NYSED requirements before, during and after test administration and ensure the integrity of the exams.

#### **Auditor Comment**

DOE's statements reveal a lack of understanding of the difference between general standards and the procedures (i.e., controls) developed to ensure that those standards are followed. NYSED is not responsible for designing and implementing specific controls; rather, each district designs controls that are unique to its respective district. The "supplemental controls" to which DOE refers are those designed to implement State requirements. DOE seems to be suggesting that the NYSED is responsible for developing specific controls for all New York State school districts. DOE would, however, readily acknowledge that controls established for schools in Erie County may not be appropriate for schools in New York City, and vice versa. Accordingly, DOE is responsible for establishing the specific controls for New York City schools to follow to ensure that the general State and Federal requirements are met.

In addition, DOE's statement that there is ongoing oversight by SCI and OSI is inaccurate. These investigative units do not perform ongoing oversight, but they do perform investigations when notified of allegations of testing improprieties. Further, as stated in the report, DOE does not have a formalized process to ensure that allegations of cheating substantiated by these agencies are shared with DOE's Office of Accountability.

#### **DOE Response**

In an April 7, 2009 preliminary draft report and at a May 4, 2009 "exit conference," audit staff informed DOE that they had focused on DOE's supplemental monitoring under an assumption that DOE was "primarily responsible for designing the controls over the [New York State] tests."

#### **Auditor Comment**

The Draft Report, the document that DOE should be responding to, does not state that DOE is primarily responsible for designing the controls, but states, "DOE is primarily responsible for coordinating yearly testing and compliance with NYSED guidelines and its own controls."

### **DOE Response**

The Comptroller never explains how controls can be characterized as inadequate when they are in all cases *above and beyond* what the law and all other governing rules, guidelines and standard practices require. [Emphasis in original.]

### **Auditor Comment**

Regarding the risk areas that the audit cites DOE for lacking adequate controls, either DOE did not identify the controls it had to address them, or the controls DOE identified were not functioning in a manner to mitigate those risks.

### **Re: Erasure Analysis**

#### **DOE Response**

The former Board abandoned the practice [of erasure analysis] because it was found to be a poor use of resources that repeatedly led to the same result the Comptroller obtained when he used the procedure in his 2007-09 audit—that no “conclusions” of manipulation or misconduct could be drawn.

#### **Auditor Comment**

DOE has yet to produce any evidence to support the claim that computerized erasure analysis was a poor use of resources. In addition, our manual erasure analysis test was conducted on a very limited basis and was not of sufficient size to be representative of the entire population. DOE argues that no “conclusions” of manipulation were drawn when the practice was used by the Board of Education; however, that is not what erasure analysis is intended to provide. Rather, it is a tool to be used in conjunction with other tools to determine whether possible manipulation has occurred.

### **Re: DOE Testing Monitors**

#### **DOE Response**

The checklist gives 7:30 a.m. as an appropriate time for monitors to arrive at the school doors in order to be admitted in time to view the unwrapping of the booklets, but monitors visiting schools with later start times (including schools where the doors do not open to permit staff to enter until after 7:30 a.m.) are sometimes admitted after 7:30 a.m. Monitors indicate on the checklist the time they entered the school building, and whether they viewed the removal of booklets from the shrink wrap.

### **Auditor Comment**

DOE's checklist requires monitors to arrive at 7:30 a.m. The purpose of arriving at 7:30 a.m. is to ensure that test materials are secured in a locked facility. When we conducted our visits to the schools, we found it necessary to arrive at 7:30 a.m. (and had no problems gaining entrance to the schools) since the majority of the schools we visited opened the shrink wrap by 8:15 a.m. In fact three of the schools we visited removed the shrink wrap from the exam booklets by 8:00 a.m., with two opening it by 7:45 a.m. Forty-seven percent of the DOE monitors arrived after 7:30 a.m., and 16 percent arrived at or after 8:00 a.m. It is difficult to believe that monitors who arrived at 8:00 a.m. or later were able to observe that test materials were secured in a locked facility and that the shrink wrap was not opened until one hour prior to the administration of the tests, especially since DOE requires that all tests begin no later than 9:15 a.m.—schools are allowed to start the exams earlier and in some cases do.

### **DOE Response**

After the auditors had conducted most of the activities and made the positive findings above, they shifted their focus again, to DOE's supplemental monitoring of a sample of testing schools. Supplemental monitoring is one of over 20 controls DOE voluntarily implements to assure the integrity of state test administration and results. When the auditors informed DOE in February 2009 that their focus had changed from a search for cheating during tests and irregularities in the handling and scoring of test materials to DOE's supplemental monitoring, DOE offered them an opportunity to observe monitoring occurring at DOE schools during March 2009 Math tests.

### **Auditor Comment**

DOE's comments show an alarming lack of understanding of audits in general and of this audit in particular. Once again, the focus of the audit was not to search for cheating. The objective was always to determine whether DOE had adequate internal controls to ensure the integrity of the assessment process. On February 25, 2009, we informed DOE of the issues regarding the weaknesses in monitoring-system controls that were identified by the audit. There was no mention of a change in audit focus. Having informed DOE of our findings, we did not accept DOE's offer to observe monitoring during March 2009 because we would have expected that the results of our review would have been discussed with all of the monitors. The monitors who might have accompanied us in March would therefore have been put on notice to ensure that all of the deficiencies that we had previously observed were corrected. Consequently, the likelihood of observing a genuine performance of the monitoring function in March would have been compromised.

### **DOE Response**

Moving beyond whether cheating on tests had occurred . . . the auditors focused on a single supplemental control DOE has adopted to assure the integrity of state testing practices.” [Emphasis in original.]

### **Auditor Comment**

DOE’s statement that the audit focused on a single control is incorrect. The audit reviewed all of the controls identified by DOE during the course of the audit, a fact that DOE well knows since it acknowledges in its response that the audit found that DOE has adequate internal controls with respect to ensuring that schools are familiar with procedures for administering the test. A careful review of the controls identified by DOE to ensure the integrity of state testing practices revealed that many of the ones related to the actual conducting of the tests are enforced at the school level. When questioned about how DOE ensures that schools are following these guidelines, DOE officials identified the testing monitors as the primary control. Accordingly, our audit included a review of this control area.

### **DOE Response**

The ten schools the Comptroller independently chose to observe during administration of the 2008 English, Math and Science tests and the seven schools he chose to observe during administration of the 2009 English tests are listed in Appendix B. The 125 elementary schools DOE randomly selected and monitored during administration of the 2008 English and Math tests . . . are listed in [Appendix] D. . . . As a comparison of those lists reveals, the Comptroller observed no school at which supplemental monitoring was taking place.

### **Auditor Comment**

What DOE fails to mention is that it subsequently created Appendix D specifically for the purpose of including it in its response to the audit. In fact, DOE did not keep track of the monitoring visits that were conducted during the administration of the 2008 ELA and Math tests.

### **Re: Analyses of 2008 Data**

### **DOE Response**

When asked at the exit conference to identify the standards or norms against which the Comptroller was measuring the adequacy of DOE’s voluntary and supplemental controls, the audit staff vaguely referenced “Internet research” and “best practices.” . . . The Comptroller has not cited any school district in the US that uses the voluntary practices he proposes or that has a package of supplemental controls that matches or surpasses the controls DOE has adopted above and beyond state requirements.

### **Auditor Comment**

DOE seems to be suggesting that safeguards are not needed if other districts have not implemented them. As stated previously, the incentive to manipulate test scores, and the risk that this may occur, has increased in recent years. It would, therefore, be in DOE's best interest to implement a strong set of internal controls to minimize the likelihood of inappropriate manipulation. Further, on May 20, 2009, the auditors e-mailed DOE a list of other states that conduct erasure analysis.

### **DOE Response**

The Comptroller conducted extensive statistical and forensic analyses to identify potential "inappropriate manipulation of test scores or cheating," including (1) review of hundreds of delivery logs, letters explaining late delivery of documents and make-up exam reports; (2) forensic (erasure) analysis of hundreds of student answer documents at schools with large score increases; (3) a detailed statistical analysis of score patterns at schools in each of the city's 32 community districts where mean scale scores increased or decreased the most; and (4) a detailed review of substantiated allegations of cheating referred to the Office of Special Investigations for the 2005-06, 2006-07, and 2007-08 school years.

After conducting these painstaking analyses, the Comptroller was "unable to draw any conclusions" of cheating, manipulation, tampering or other irregularities (p. 16). Nevertheless, the Comptroller opines that DOE's internal controls are inadequate.

### **Auditor Comment**

It is unclear whether DOE does not understand the report or is intentionally misstating the audit's findings. As we clearly state in the report, "we were unable to draw any conclusions from our *attempted* data-trend analysis *due to the very limited information available.*" (Emphasis added.) There was not enough data available to conduct a true trend analysis of late deliveries, make-up exams, or score patterns. A minimum of three to four years of data is needed in order to ascertain whether a trend exists. DOE maintained only one year's worth of data. With regard to erasure analysis, the report clearly states that it was conducted on a limited basis. Lacking the software usually employed to conduct such an analysis, we performed a non-statistical analysis, selecting only 15 schools, and only one fourth grade class within each school. Due to the limited nature of our analysis, our results can by no means be used by DOE as evidence that erasure analysis conducted on a larger scale would yield the same results.

We are troubled by DOE's attempt to change the meaning of our statement by deleting the qualifying statement highlighted above. DOE is attempting to manipulate our report to present "findings" that the report does not make.



**Re: Monitoring Checklists**

**DOE Response**

The Comptroller is correct that a few items on the 2008 list [monitoring checklist] were unnecessary, but those items were removed or re-written during administration of the 2009 state tests—as the auditors would have seen, if they had accepted DOE’s invitation to observe monitoring taking place during 2009 Math tests.

**Auditor Comment**

DOE’s claim that items on the checklist were removed or rewritten during the administration of the 2009 state tests is erroneous and implies that the revisions were not based on our audit findings. However, the evidence appears to indicate otherwise. We informed DOE of our concerns regarding the questions on the checklist on February 25, 2009. At that time, DOE officials provided no evidence that the checklist had been modified. In fact, it was not until the exit conference on May 4, 2009, that DOE officials first informed us that the checklist had been revised. (However, we were not provided a copy until May 8.)

**Re: Sharing of Substantiated Allegations of Cheating**

**DOE Response**

DAAR is aware of allegations during SCI and OSI investigations. When an allegation of a testing violation is made, DAAR often is contacted first and in some cases refers the matter to SCI. Once SCI or OSI undertakes an investigation, it works closely with a DAAR specialist to investigate the allegations. The DAAR specialist has deep knowledge of testing practices, provides continuous feedback to her DAAR colleagues on potential risk areas with respect to testing misconduct and proposes changes in testing practices to improve security processes and procedures in city schools.

**Auditor Comment**

According to DOE officials, testing improprieties or irregularities may be reported through a number of channels including SCI, OSI, ISC, DAAR, the Assessment Implementation Director (AID), NYSED, the DOE help desk, and scoring sites. In addition, DOE is required to report allegations of misconduct to the school principal, SCI, the AID at the ISC, and to NYSED. However, DOE does not maintain a central log of all allegations reported. Without a central log or formal recordkeeping process, DOE cannot demonstrate that all allegations of testing improprieties were recorded, reported to the required parties, including DAAR, and investigated. As a result, DOE has limited assurance that DAAR is aware of all allegations and potential risk areas and would be unable to propose changes in test practices to improve security.

**Re: Inadequate Preventive and Detective Controls**

**DOE Response**

DOE asked the auditors to share the specific sources, online or otherwise, from which they derived their standards or norms, or at least to identify one or more school districts that use the “best practices” they recommend. The auditors provided no such citations, and instead acknowledged that their assessments of adequacy are based entirely on their own, lay opinions and beliefs about whether DOE’s voluntary controls were adequate—absent any background or expertise on their part in the detection or prevention of cheating or manipulation in test administration. The Comptroller has not cited any school district in the US that uses the voluntary practices he proposes or that has a package of supplemental controls that matches or surpasses the controls DOE has adopted.

**Auditor Comment**

Our assessment is based on the simple fact that very few of the controls identified by DOE are designed to prevent and detect irregularities, and those that are either are not formalized or are incomplete. For example, DOE requires that the score sheets for Day One of the Math and ELA exams are removed from the schools and forwarded to the scan centers. However, there is no similar control in place for Day Two and Day Three exams, nor for the makeup exams.

Additionally, DOE’s statement that we did not provide the names of school districts that perform some of the analyses we discuss is incorrect. We provided DOE with the names of some of the states whose districts conducted these analyses. However, it is not our role to research the methods used elsewhere that may be transferred to DOE schools. That responsibility belongs to DOE. DOE’s request for the results of our research appears to be an attempt to find fault with any methodologies we recommend rather than use the underlying basis of our recommendation, namely to find or develop methodologies that could be used effectively in New York City schools.

**DOE Response**

The Comptroller’s belief that there is a risk of improper manipulation of Day Two and Three materials is based on a single substantiated case over three years and well over a million Day Two and Day Three test booklets. That isolated instance is no basis for changing procedures that follow NYSED requirements and have worked exceedingly well to avoid manipulation across millions and millions of student tests.

**Auditor Comment**

Our belief that there is a risk of improper manipulation of Day Two and Three materials is based on the fact that these tests are maintained at the schools until the completion of the

make-up exams—up to seven days from the initial administration of the exam. Allowing the schools to maintain the answer documents for this period of time increases the risk of inappropriate manipulation.

Based on DOE’s comments, it appears that the agency believes there is virtually no manipulation of Day Two and Day Three tests because there was only one substantiated case of improper manipulation. We do not necessarily agree with this position. It may very well be that no other instances were detected because DOE does not have adequate detection tools in place. For example, although DOE requires schools to keep these exams locked and secured, DOE has no procedure or control in place to verify that it is being done.

**Re: Audit Protocol**

**DOE Response**

When the Comptroller’s staff produced a preliminary draft report acknowledging that their erasure analysis revealed no evidence of cheating or manipulation, the Comptroller violated longstanding protocol by pulling new allegations out of the hat at the last minute when timely response was difficult.

**Auditor Comment**

DOE is disingenuous in this response. This information was added to the draft report *at the request of DOE officials* at the exit conference. DOE’s comments reveal a lack of understanding of the purpose of the preliminary draft report. We are not required to issue a preliminary draft report, but do so as a courtesy to help facilitate the discussion of audit results at the exit conference. Based on these discussions, we frequently make revisions that are reflected in the official draft report. Accordingly, adding additional information to the draft report regarding erasure analysis was hardly a violation of protocol but, in fact, an accommodation to DOE officials at the exit conference.

**Conclusion**

Overall, after carefully reviewing DOE’s arguments, we found them to be without merit. Accordingly, we stand by our findings.



June 16, 2009

James S. Liebman  
Chief Accountability Officer  
JLiebman@schools.nyc.gov

52 Chambers Street  
New York, NY 10007

+1 212 374 4941 tel

Mr. John Graham  
Deputy Comptroller  
Audits, Accountancy and Contracts  
The City of New York Office of the Comptroller  
1 Centre Street  
New York, NY 10007-2341

Dear Mr. Graham:

This letter, together with the enclosed Response to Findings and Recommendations, and Appendices A-H, constitute the New York City Department of Education's response to the New York City Office of the Comptroller's June 2, 2009 draft report on the Administration of New York State Standardized Tests by the New York City Department of Education (Audit Report # MD08-102A).

Sincerely,

A handwritten signature in black ink, appearing to read "James S. Liebman". The signature is written in a cursive style with a long, sweeping underline that extends to the left.

James S. Liebman  
Chief Accountability Officer

# NEW YORK CITY DEPARTMENT OF EDUCATION'S RESPONSE TO COMPTROLLER'S AUDIT REPORT MD08-102A ON ADMINISTRATION OF NEW YORK STATE STANDARDIZED TESTS

June 16, 2009

## EXECUTIVE SUMMARY

Each year, the New York City Department of Education (DOE) administers New York State standardized tests to students in grades 3 through 8 in accordance with requirements set by the New York State Education Department (NYSED).

In December 2007, City Comptroller William C. Thompson, Jr. (Comptroller) began an audit of DOE's administration of the 2008 English Language Arts (English) and Math tests. After a lengthy and exhaustive audit—which expanded once to include test handling, delivery, scanning, scoring and reporting, again to include the 2008 Science and 2009 English exams, and again to include painstaking statistical and forensic analyses of a large sample of test answers—the Comptroller found that **city personnel are well-trained in state and local testing requirements and that they administer tests in compliance with those requirements. *No instances of cheating, manipulation, tampering, misconduct, or other irregularities were found.***

Eighteen months after the audit began, the Comptroller is finally releasing his report, which ignores his auditors' entirely positive findings about the integrity of the tests and their results and makes a series of trivial debater's points, replete with clear mistakes and loud demands for changes DOE already made months ago. Instead of confirming that his auditors found no problems on more than a million student exams and proposing modest process improvements, the report complains about the way DOE *used to* implement a *single* procedural control that DOE has *voluntarily* adopted—above and beyond anything federal and state law require and any other district in the country uses—which the Comptroller *never actually observed in action and badly misrepresents*. And instead of celebrating test score gains that the audit has exhaustively validated across two years and hundreds of thousands of children, the report levels last-minute, irresponsible accusations against two fourth graders for making—and correcting—the most common of “bubbling” mistakes (each obviously got off by a row or two in moving answers from test booklet to bubble sheet).

That the Comptroller has had to struggle so hard to find something to complain about in a comprehensive 18-month audit is a testament to the accuracy and integrity of test results showing years of impressive learning gains by New York City children and their teachers. One can only wonder why the Comptroller would go to such lengths to find fault where clearly none exists.

### ***1. The Comptroller Found No Instances of Cheating During the Administration of the 2008 State English, Math and Science and the 2009 State English Tests***

In December 2007, the Comptroller notified DOE of his intent to audit DOE's administration of state English and Math tests during the 2007-08 school year to consider the integrity of the examination process and look for evidence of “cheating.” Although the Comptroller's formal audit notification was limited to test administration during the winter 2008 English and Math tests, the Comptroller subsequently and without notification expanded the scope of the audit to include the

April 2008 state Science tests, and expanded it again to include the 2009 state English tests. During the audit, members of the Comptroller's staff made unannounced observations of the administration of each of the four tests at a number of city schools selected by the Comptroller.

After an 18-month search for evidence of cheating on four separate tests across two school years and three subjects, generating approximately 1.8 million answer documents, the Comptroller found *no evidence of cheating*: "Our ... review of the data and documentation collected by DOE for the 2007-2008 English and Math tests and our observations conducted at the sampled schools on the day of testing *did not reveal any instances of cheating*" (Draft Audit Report, p. 1-2). "DOE has *adequate internal controls* with respect to ensuring that schools are familiar with established procedures when administering the New York State standardized tests at elementary schools" (p. 1). "[T]he schools that we visited generally *complied with the State testing guidelines, DOE Handbook, and testing memoranda*" (p. 1). The Draft Report likewise found *no instances of cheating* on the 2008 Science and 2009 English tests.

**2. The Comptroller Found No Manipulation, Tampering or Irregularities in the Handling, Storing, Scoring and Reporting of the 2008 and 2009 English, Math and Science Tests**

After months spent searching for cheating during administration of state tests, the Comptroller expanded the audit again to search for evidence of "inappropriate manipulation," "tampering" or "irregularities" in DOE's handling, delivery, scanning, scoring and reporting of state tests and their results. After dozens of interviews and on-site visits, meticulous statistical and forensic (including erasure) analyses of test results and answer documents and a careful review of three years of allegations received by the Office of Special Investigations of misconduct in test administration, the Comptroller was *"unable to draw any conclusions" that misconduct had occurred* (p. 16).

The Comptroller's findings based on his observations and statistical and forensic analyses of test administration, handling, scoring, and reporting are set out below:

Analysis Conducted	Findings
"physical observations of the [2008 and 2009 English and 2008 Math] test-taking process ... to test compliance with the controls that were described to us in the interviews or that were contained in policies and procedures" (p. 5)	"did not reveal any instances of cheating" (p. 9); schools "complied with the State testing guidelines, DOE Handbook, and testing memoranda" (p. 1)
"review of the administration of the [2008] Science test ... to observe compliance with the controls described to us in the interviews or contained in policies and procedures" (p. 6)	"did not reveal any instances of cheating" (p. 9); schools "complied with the State testing guidelines, DOE Handbook, and testing memoranda" (p. 1)
"observe ... testing classrooms for the duration of the exam" (p. 11)	"did not reveal any instances of cheating" (p. 9); schools "complied with the State testing guidelines, DOE Handbook, and testing memoranda" (p. 1)
"visited the Queens ISC [twice], and observed the delivery of completed tests by Queens elementary schools and the pick-up of the tests by the courier service for delivery to the Scan Center" (p.6)	found no instances of cheating, manipulation, tampering or other irregularities
"visited the Scan Center, observed the delivery of tests by the courier service, and conducted a walkthrough of the operations of the Scan Center" (p. 6)	found no instances of cheating, manipulation, tampering or other irregularities

Analysis Conducted	Findings
“unannounced visits to the Manhattan scoring site during the scoring of ELA and Math tests” (p. 6)	found no instances of cheating, manipulation, tampering or other irregularities
“calculated the differences between the mean scale scores for [2006-2007 and 2007-2008] . . . English and Math tests . . . [;] identified three schools within each district that had the greatest increase in mean scale scores for each grade [and] determined whether there were any substantiated allegations of cheating or a large number of make-up exams” (p. 7)	“unable to draw any conclusions” that cheating, manipulation, tampering or other irregularities occurred (p.16)
“determined for these schools whether the multiple choice-portion of the tests were delivered late to the Scan Center in the 2007-2008 school year” (p. 7)	“unable to draw any conclusions” that cheating, manipulation, tampering or other irregularities occurred (p.16)
“for both the English and Math tests, . . . identified the three schools within each district that had the greatest decrease in mean scale score for each grade . . . [and] determined whether there were substantiated allegations of cheating in the 2006-2007 school year [and] whether a DOE monitor visited the schools during the administration of the tests that took place during the 2007-2008 school year” (p. 7)	“unable to draw any conclusions” that cheating, manipulation, tampering or other irregularities occurred (p.16)
“erasure analysis of the test answer documents from one fourth-grade class from each of . . . 15 sampled schools” by “manually review[ing] 336 student English answer documents” (p. 14-15 & n.5)	“unable to draw any conclusions” that cheating, manipulation, tampering or other irregularities occurred (p.16)

**3. *The Comptroller Finds Fault with Only 1 of 20+ Controls DOE Has Voluntarily Adopted, Above and Beyond State Requirements, and he Badly Misunderstands that Procedure, which his Auditors Refused to Observe in Action and DOE Has Already Improved***

Fourteen months into the audit, the Comptroller informed DOE that he had expanded the audit again. Moving beyond whether *cheating on tests* had occurred (after *none was found*), and beyond whether *irregularities in testing procedure* had occurred (after *none were found*), the auditors focused on a single supplemental control DOE has adopted to assure the integrity of state testing practices. As is outlined in Appendix A, DOE has adopted over 20 supplemental controls to assure the integrity of test results. DOE undertakes each control voluntarily, above and beyond state requirements, and supplementary to ongoing oversight both by NYSED and by two investigative agencies (the Office of the Special Commissioner of Investigation and the Office of Special Investigations). The single supplemental control the Comptroller looked at is DOE’s unannounced monitoring of a sample of schools (hereinafter “supplemental monitoring”), which DOE voluntarily conducts to augment New York State’s own unannounced monitoring.

When the auditors first identified this new focus, in February 2009, DOE invited them to observe supplemental monitoring taking place during the March Math tests. DOE made the offer because supplemental monitoring occurs at only a random sample of the City’s 1,100 elementary and middle schools, and none of those schools was among the 17 the auditors had visited to observe test administration. The auditors declined the invitation and, as a result, have never observed DOE’s supplemental monitoring.

In an April 7, 2009 preliminary draft report and at a May 4, 2009 “exit conference,” audit staff informed DOE that they had focused on DOE’s supplemental monitoring under an assumption that

DOE was “primarily responsible for designing the controls over the [New York State] tests.” In fact, DOE is not primarily responsible for setting testing controls and procedures. Federal and state law give that responsibility to NYSED. The single additional control on which the Comptroller focused is entirely voluntary on DOE’s part.

When DOE so informed the Comptroller’s staff at the exit conference, and asked what the audit had revealed about DOE compliance with NYSED requirements, what rules, guidelines, standards or accepted procedures apart from state law they had used to evaluate DOE’s supplemental controls and whether those controls failed to meet a standard of practice in any other school district in the state or nation, the audit staff acknowledged that (1) despite receiving NYSED testing procedure documentation from DOE during the course of the audit fieldwork, they were not aware of NYSED’s primary responsibility for defining state test security procedures and requirements, (2) had not understood there to be a difference, nor had they distinguished between, NYSED’s mandatory controls and DOE’s voluntary controls,<sup>1</sup> (3) had not spoken to anyone at NYSED about its controls or its evaluation of controls in place in New York City and (4) would have to “get back” to DOE about test control standards beyond state law—or practices in other districts—that the audit team had used in their audit and were more demanding than DOE’s controls.

In his Draft Report, the Comptroller has removed language ascribing “primary responsibility” to DOE for establishing test security controls and finds that DOE and its schools “complied with the State testing guidelines” (p. 1). He continues, however, to opine that one of DOE’s voluntary controls, supplemental monitoring, is not “adequate.” This assertion is flawed in several respects:

- Despite numerous requests that he do so, the Comptroller points to no laws, regulations, standards, guidelines, accepted procedures or even individual practices of another school district among the 14,000 in the US that DOE’s supplemental practices fail to satisfy or surpass. The only applicable rules the Comptroller identifies are NYSED’s, and he acknowledges that DOE “complies” with those rules (p.1). The Comptroller never explains how controls can be characterized as inadequate when they are in all cases *above and beyond* what the law and all other governing rules, guidelines and standard practices require.
  - The auditors never observed supplemental monitoring in action. DOE offered them the chance to do so, but they declined. As a result, they repeatedly misdescribe the practice.
  - The Draft Audit Report persistently uses the present tense to describe aspects of DOE’s supplemental monitoring that DOE improved months ago, in time for the 2009 testing cycle—as the auditors would have known if they had accepted DOE’s offer to observe the practice.
- 4. In Advocating a Type of Forensic Analysis his own Auditors Used at Great Expense but with No Results, and that the Former Board of Education Abandoned in 2001 Because it Was Costly and Had No Results, the Comptroller Egregiously Misstates the Facts and Falsely Accuses Two Fourth Graders***

Finally, the Comptroller’s Draft Report criticizes DOE for failing to resurrect a practice (systematic erasure analysis) that the former Board of Education eliminated in 2001—when the

---

<sup>1</sup> The Comptroller remains confused. His Draft Audit Report (p. 13) refers to the rule barring removal of shrink wrap from test booklets until an hour before the exam as a “State guideline.” In fact, this is a voluntary *DOE* control put in place to augment the state rule allowing shrink wrap to be removed at any time on “the *date* on which” tests are given.



Comptroller himself was the Board's President. The former Board abandoned the practice because it was found to be a poor use of resources that repeatedly led to the same result the Comptroller obtained when he used the procedure in his 2007-09 audit—that no “conclusions” of manipulation or misconduct could be drawn (p. 16). Most egregiously, the 2 out of 336 answers sheets the Comptroller analyzed that he suggests may reveal an “excessive pattern of erasures” (p. 15) are both obvious cases—confirmed by a testing expert (see Appendix H)—where the student got off by one or two rows in bubbling-in answers on the answer sheet and simply moved the prior sequence of answers into the right rows after realizing the mistake.

\* \* \* \* \*

The attached response describes DOE test administration and supplemental controls and the Comptroller's audit of both. It does so to call attention to two peculiarities in the Comptroller's draft report. The first is its de-emphasis of the auditors' favorable findings on the absence of cheating, manipulation, tampering or any other irregularities; DOE compliance with the governing federal and state controls; and the integrity of DOE test administration, delivery, handling, scoring, scanning and reporting and DOE's numerous supplemental controls. The second is the Draft Report's oddly overreaching efforts to manufacture problems where, clearly, none exist.

## **NEW YORK CITY DEPARTMENT OF EDUCATION'S RESPONSE TO COMPTROLLER'S AUDIT REPORT MD08-102A ON ADMINISTRATION OF NEW YORK STATE STANDARDIZED TESTS**

June 16, 2009

### **I. FEDERAL AND STATE TESTING REQUIREMENTS**

The federal No Child Left Behind Act requires states to administer annual tests to measure student proficiency in English, Math and other subjects and holds states responsible for establishing and enforcing rules and procedures to ensure that the tests are administered and scored in a secure and reliable manner.<sup>2</sup> In compliance with these requirements, NYSED tests all eligible students in grades 3-8 in English and Math and tests 4<sup>th</sup>, 5<sup>th</sup> and 8<sup>th</sup> graders in Science and/or Social Studies. NYSED sets the testing calendar—requiring that the English test be administered in January, the Math test in March, and the Science and Social Studies tests in the fall or spring—and it has established a comprehensive and exacting set of requirements that all schools in the state must observe when administering and scoring tests. In New York City each year, English and Math test administrations occur in tens of thousands of classrooms across more than a thousand schools in grades 3-8, resulting in roughly 1.8 million separate test answer sheets. Additional administrations occur in Science and Social Studies.

The federal and state governments use test results to evaluate schools and impose serious consequences on those that are not making “adequate yearly progress.” New York City also uses test results on its annual Progress Reports to measure how well schools help students make progress in mastering state learning standards, with bonuses for high performing schools and consequences for chronically low performing schools. Principals and teachers use test results to make program decisions for students, and to determine whether to promote students to the next grade. Parents use federal, state and city accountability outcomes based on test scores to choose schools for their children, and communities use them to demand improved results.

### **II. THE COMPTROLLER'S COMPREHENSIVE AUDIT OF THE 2008 ENGLISH, MATH AND SCIENCE TESTS AND 2009 ENGLISH TESTS**

In December 2007, the Comptroller announced his intent to audit DOE's administration of New York State tests during the 2007-08 school year.

The Comptroller's painstaking 18-month audit required the participation of two or three auditors throughout the audit period and additional staff during periods when fieldwork was being conducted. The Comptroller has declined DOE's request to disclose the number of people and amount of resources dedicated to supporting his field, documentary, statistical and forensic review.

---

<sup>2</sup>The NCLB No Child Left Behind Act requires NYSED to “set academic standards for what every child should know” English and Math and “test students in grades 3–8 ... on what they know in ELA and Math,” and makes NYSED responsible for “ensuring the continued validity and reliability of State assessments . . . and developing multiple measures to increase the reliability and validity of State assessment systems.” 20 U.S.C. §§ 7301(1), 7301(2)(C), 7301(2)(D); <http://www.cmssc.nysed.gov/deputy/nclb/parents/facts/eng/eng-overview.html>.

The auditors' 2008 fieldwork, focused on English, Math and Science tests, included:

- *Unannounced Visits to Schools.* Auditors selected and visited 10 schools in January, March and April 2008 to observe administration of English, Math and Science exams.
- *Unannounced Visits to Scoring Site.* Auditors conducted two unannounced visits to a scoring site to observe scoring of written portions of the English and Math exams.
- *Visits to Integrated Service Center and Scan Center.* Auditors visited an Integrated Service Center and the Scan Center multiple times to observe security procedures associated with delivery and handling of completed answer documents.

As revealed by the Draft Audit Report and the auditors' document requests, the Comptroller's documentary, statistical and forensic analysis of the 2008 English and Math tests included:

- *Review of NYSED and DOE test procedure manuals.* Auditors reviewed manuals and documentation produced by NYSED and DOE regarding testing procedures, including the New York State Testing Program *School Administrator's Manuals* for English and Math, *Teacher Directions* for English and Math, *NYC Test Administration Handbook*, *NYC Test Administration Memorandums*, *Directions for Scribing*, *NYC Test Security Monitoring Checklists*, *Class Header and Answer Documents*, *Test Material Security Form*, and materials pertaining to DOE promotion policy.
- *Review of Delivery Logs, Letters Explaining Late Delivery, and Make-up Exam Reports.* Auditors reviewed hundreds of test delivery documents and make-up exam reports to identify abnormal delivery schedules and numbers of make-up exams that might indicate manipulation and to verify completion of all documentation. These documents were compared to school results in order to identify patterns that might indicate testing irregularities.
- *Forensic (Erasure) Analysis.* In search of possible evidence of cheating or misconduct, auditors reviewed 336 student answer documents and scores from 15 4<sup>th</sup> grade classrooms at 15 separate schools to identify patterns of changed answers that might indicate manipulation. The auditors chose schools and classes for erasure analysis by examining score patterns and selecting schools with the largest improvements in test scores.
- *Review of Abnormal Score Patterns.* Auditors analyzed students' mean scale scores by school to identify three schools in each of the 32 community districts in the city where students' mean scale scores increased or decreased the most in a year and examined whether the Office of Special Investigations (OSI) had substantiated any allegations of cheating in those schools.
- *Review of Allegations.* Auditors conducted a detailed review of substantiated allegations of cheating that were referred to the Office of Special Investigations for the 2005-06, 2006-07, and 2007-08 school years. Auditors also reviewed DAAR records connected to investigations of 2007-08 testing allegations relating to the NYS English and Math exams and interviewed the DAAR employee who assists OSI in conducting the investigations.

The Comptroller's 2009 fieldwork, focused on English tests, included:

- *Unannounced Visits to Schools.* Auditors observed administration of English tests at 7 additional schools in January 2009. Again, the auditors appeared unannounced at schools chosen independently by the Comptroller.
- *Visits to Integrated Service Centers and Scan Center.* Auditors again visited an Integrated Service Center and the Scan Center to observe security procedures associated with delivery and handling completed answer documents.

### **III. THE COMPTROLLER'S CONSISTENTLY POSITIVE FINDINGS ON DOE'S ADMINISTRATION, DELIVERY, HANDLING, SCANNING, SCORING AND REPORTING OF THE 2008 AND 2009 ENGLISH, MATH AND SCIENCE TESTS**

After conducting the extensive fieldwork, interviews, document review, statistical analysis and other actions listed above, the Comptroller found *no problems* with the administration of the 2008 English, Math and Science tests and the 2009 English test and *no evidence of cheating, tampering, manipulation or other irregularities* in any of the four test administrations. *All testing practices were found to comply with the governing state requirements.* Testing personnel were found to be *well-trained in state requirements and city guidelines* relating to test administration<sup>3</sup> and to have *administered the tests in compliance with those requirements.*<sup>4</sup> *No "instances of cheating" were found* (p. 2),<sup>5</sup> and *no "conclusions"* (p. 16) of fraud, manipulation or misconduct were made or allegations referred to the Special Commissioner of Investigation for further review.

The expanded audit also scrutinized handling and delivery of test materials and answer documents that move between schools, scanning center and scoring sites and the scanning and hand-scoring of hundreds of thousands of test answer documents. These investigations covered three separate test administrations in January and March 2008 and January 2009. Again, the Comptroller found *no manipulation, misconduct, tampering or other irregularities.*

### **IV. THE COMPTROLLER'S AUDIT OF DOE'S SUPPLEMENTAL MONITORING AND LACK OF STANDARDS FOR EVALUATING DOE'S VOLUNTARY CONTROLS**

After the auditors had conducted most of the activities and made the positive findings above, they shifted their focus again, to DOE's supplemental monitoring of a sample of testing schools. Supplemental monitoring is one of over 20 controls DOE voluntarily implements to assure the integrity of state test administration and results. When the auditors informed DOE in February 2009 that their focus had changed from a search for cheating during tests and irregularities in the handling and scoring of test materials to DOE's supplemental monitoring, DOE offered them an opportunity to observe monitoring occurring at DOE schools during March 2009 Math tests. DOE pointed out that it only monitors a sample of schools, and that none of the 17 schools the auditors had chosen to visit in January, March and April 2008 or January 2009 was monitored during the

---

<sup>3</sup> "DOE has adequate internal controls with respect to ensuring that schools are familiar with established procedures when administering the New York State standardized tests at elementary schools" (p. 1).

<sup>4</sup> "[T]he schools that we visited complied with the State guidelines and the guidelines outlined in the Handbook" (p. 1). [*Test Administration Handbook: Elementary – Middle Schools*].

<sup>5</sup> "Our own review of the data and documentation collected by DOE for the 2007-08 [English] and Math tests and our observations conducted at the sampled schools on the day of testing did not reveal any instances of cheating" (p. 1-2).

test administrations the auditors observed. The auditors declined. Nonetheless, the Draft Report devotes most of its findings to this single supplemental control.

When asked at the exit conference to identify the standards or norms against which the Comptroller was measuring the adequacy of DOE's voluntary and supplemental controls, the audit staff vaguely referenced "Internet research" and "best practices." DOE asked the auditors to share the specific sources, online or otherwise, from which they derived their standards or norms, or at least to identify one or more school districts that use the "best practices" they recommend. The auditors provided no such citations, and instead acknowledged that their assessments of adequacy are based entirely on their own, lay opinions and beliefs about whether DOE's voluntary controls were adequate—absent any background or expertise on their part in the detection or prevention of cheating or manipulation in test administration. The Comptroller has not cited any school district in the US that uses the voluntary practices he proposes or that has a package of supplemental controls that matches or surpasses the controls DOE has adopted above and beyond state requirements.<sup>6</sup>

## V. DOE'S SUPPLEMENTAL TEST SECURITY PROCESSES

NYSED administers annual English, Math and Science tests and is responsible for defining the processes and procedures to be used in administering and ensuring the security of the tests. (See p. 1, fn. 2 above.) The Office of the Special Commissioner of Investigation (SCI) and Office of Special Investigations (OSI) have ongoing responsibility to investigate allegations, and recommend sanctions, for misconduct in test administration.

As additional precautions, DOE voluntarily implements a set of procedures, not required by federal or state law, to verify schools' adherence to NYSED requirements before, during and after test administration and ensure the integrity of the exams. Appendix A lists NYSED's mandated practices and DOE's voluntary controls. DOE's supplemental controls have more than 20 components, one of which is the unannounced monitoring of a random sample of testing schools.

NYSED itself makes unannounced visits to schools throughout the state to verify that exam materials are properly stored and that proper test administration procedures are used. NYSED does not require school districts to conduct their own spot checks and monitoring, but DOE does so. DOE's supplemental monitoring consists of unannounced visits to review the entire testing process at visited schools to ensure adherence to NYSED-defined testing protocols and to serve as a deterrent for all schools.

DOE supplemental monitoring takes place as follows:

- Several weeks before tests are administered, the Office of Assessment identifies volunteers from a broad pool of central employees who are willing to participate in monitoring during an upcoming test administration, assuming their ongoing central responsibilities permit them to participate on the days monitors are needed. Using DOE volunteers who are able to take time away from their regular duties in order to carry out this function helps DOE conduct supplemental monitoring at little or no additional cost. (Because NYSED doesn't require supplemental monitoring, it doesn't reimburse DOE for its costs.) Volunteers indicate their

---

<sup>6</sup> For example, according to information the audit staff reported back to DOE, erasure analysis is a matter of *state*, not school district, practice, and only a few states use it.

tentative availability and preferred geographic area. (Monitoring begins first thing in the morning, so an effort is made to assign monitors based on their residential locations.)

- Based on the number of potential volunteers and their locations, a subset of schools is randomly selected, a tentative monitoring schedule is developed and tentative assignments are made—approximately two weeks before the relevant testing dates. Because volunteers' possible participation is solicited weeks in advance and is premised on their ability to take time away from their ongoing DOE duties on the day actually requested, and because schedules projected weeks in advance change, it is assumed that a number of prospective monitors will not be able to accept their tentative assignment. The schools actually monitored are determined based on the volunteers who are available to accept a proposed assignment.
- During the unannounced visit, monitors discuss procedures with school staff to assure the school adheres to NYSED protocols and observe test administration at the school. Volunteers use interviews and observation to obtain information about procedures in all classrooms school-wide, and observe a subset of the classrooms through windows. (NYSED regulations forbid individuals other than teachers serving as test proctors to enter classrooms during testing.)
- Monitors use a monitoring checklist (a copy is in Appendix C) to remind them of the security processes and protocols that schools are required to implement before, during and after test administration. The checklist prompts monitors to confirm, for example, that test booklets are stored in secure locations and that shrink wrap is not opened or removed until shortly before the testing begins; that all materials that might improperly aid students have been covered or erased; that students are seated so they cannot see other students' answer documents; and that teachers sign a receipt and class roster when they pick up and return exams.
- One of the test administration procedures monitors observe is the unwrapping of the shrink-wrapped test booklets. The unwrapping occurs on the morning of the tests and, in accordance with DOE's supplemental policies, no more than one hour before the test begins ("the hour period").<sup>7</sup> Depending upon when schools open, which varies throughout the city, and depending upon how much of the hour period schools choose to use, the removal of the shrink-wrap can occur as early as 7:30 a.m. and as late as 9:00 a.m. It usually occurs between 8:00 and 9:00 a.m. The checklist gives 7:30 a.m. as an appropriate time for monitors to arrive at the school doors in order to be admitted in time to view the unwrapping of the booklets, but monitors visiting schools with later start times (including schools where the doors do not open to permit staff to enter until after 7:30 a.m.) are sometimes admitted after 7:30 a.m. Monitors indicate on the checklist the time they entered the school building, and whether they viewed the removal of booklets from the shrink wrap.
- Monitors are instructed to notify DAAR immediately and directly if any irregularities occur during the visit. When a monitor notifies DAAR of a possible infraction, follow-up with the school occurs through appropriate channels, including escalation to the Office of the Special Commissioner of Investigations (SCI), a NYC office operating independently from the DOE.
- Monitors are asked to submit a completed monitoring checklist to DAAR immediately after monitoring, and Scan Center staff review each form and take any action that is warranted.

---

<sup>7</sup> NYSED requires that shrink wrap remain on the exams until "the dates on which they will be administered"; DOE requires that the shrink wrap remain until one hour prior to testing.

Although the checklist is primarily a reference for use during the monitoring visit, and although direct notification of DAAR is the means by which monitors communicate testing irregularities in the rare cases in which they are observed, Scan Center personnel review the monitoring checklist as a back-up mechanism for spotting problems.

DOE continually identifies ways to enhance test security. In 2009, for example, it improved its supplemental monitoring forms and scheduling process in the following respects:

- revising and clarifying items on the checklist.
- revising monitoring instructions to clarify expectations, including the one-hour rule on removing the shrink wrap.
- improving the scheduling process, including reconfirming scheduled visits preceding test administration to continuously update the schedule in the days leading up to each assigned visit
- confirming completed visits and following up to assure that monitoring checklists are filed for all monitoring visits completed.

During the January and March 2008 administrations of the English and Math tests at elementary schools, 269 monitoring visits at a randomly selected subset of 208 schools were tentatively scheduled. After adjusting for volunteers' schedules, 183 confirmed monitoring visits occurred, and 147 monitoring checklists (80%) were filed at the Scan Center. Of the 147 forms received, 144 (98%) note that shrink-wrap was removed from test booklets at the appropriate time (the remaining three forms note that the monitor was not present to verify). 136 (93%) indicate that the monitor was admitted to the school at or before 8:00 a.m. on the morning testing took place.

During the January and March 2009 English and Math test administrations, after scheduling practices were improved, 223 unannounced monitoring visits at a randomly selected subset of 114 schools were conducted, and monitoring checklists for all 223 of those visits (100%) were filed. 98% of monitors indicated that shrink-wrap was removed from the test booklets at the appropriate time. (The remaining two forms indicate the monitor was not present to verify.)

## **VI. THE COMPTROLLER'S REFUSAL TO OBSERVE SUPPLEMENTAL MONITORING TAKING PLACE**

The ten schools the Comptroller independently chose to observe during administration of the 2008 English, Math and Science tests and the seven schools he chose to observe during administration of the 2009 English tests are listed in Appendix B. The 125 elementary schools DOE randomly selected and monitored during administration of the 2008 English and Math tests, and the 75 elementary schools DOE randomly selected and monitored during administration of the 2009 English and Math tests, are listed in Appendices D and E. As a comparison of those lists reveals, the Comptroller observed no school at which supplemental monitoring was taking place.

On March 5, 2009, DOE offered the Comptroller the opportunity to view supplemental monitoring taking place during 2009 Math test administration. The Comptroller declined the invitation, stating that no more field work would be conducted.

During the Comptroller's winter 2008 observation of the eight schools chosen by his auditors, DOE's Deputy Executive Director of Assessment Support and Summative Assessment assigned six DOE employees to accompany the auditors on their eight visits, to assist the auditors, relieve schools of any extra burden the auditors created and assure that the auditors observed state test administration restrictions (e.g., not entering classrooms during testing). Some of these audit aides had previously served as DOE monitors, and for six of the eight audit visits, audit aides used the checklist to record activities they observed while assisting and accompanying the auditors. The aides were not assigned nor expected to conduct supplemental monitoring at the eight schools, nor did the auditors ever inform DOE that they expected to observe monitoring during their school visits at these eight or any other schools. In particular:

- As noted in the Comptroller's Draft Audit Report (p. 10), the eight schools involved were not on the tentative or final list of schools DOE had randomly selected for supplemental monitoring during the auditors' observations (see Appendices D and F).
- An explanation of the audit aides' role was sent to the Comptroller's auditors by email on February 28, 2008, prior to the March 2008 Math exam, with cc's to the audit aides (see Appendix G). The email informs the Comptroller's auditors that the aides "will meet you at the principal's office at the schools" to "assist you [the auditors]" in carrying out the Comptroller's auditing and monitoring activities. There were no instructions given to the audit aides to monitor test administration apart from activities the Comptroller's auditors chose to observe.
- The communications and timing of communication with the six audit aides were different from those used with supplemental monitors. Potential monitors are asked to volunteer for supplemental monitoring several weeks before the test dates and are informed about their proposed assignment about two weeks before the test date. By contrast, the audit aides were asked to accompany the auditors just a few days before the audit visits.
- In the e-mails in which monitoring assignments are distributed, the monitoring checklist is always attached. The checklist was not attached to the email sent to the audit aides.
- In the winter 2008 period when the auditors' eight school observations in question occurred, the auditors were focused entirely on test *administration*. It was not until many months later that their focus switched to test *monitoring*. As a result, during the winter 2008, the auditors never asked to observe schools being monitored, and were never given a list of monitored schools to choose from, nor did they by happenstance choose monitored schools. As noted, when the auditors first told DOE of their focus on supplemental monitoring, in February 2009, and when DOE as a result offered them the chance to view supplemental monitoring taking place during the March 2009 Math tests, the auditors declined the offer. The reason the auditors gave for declining the offer was *not* that they had already viewed supplemental monitoring but instead that they had decided against conducting additional field work.
- In August 2008, the Comptroller specifically requested forms filed by DOE personnel related to testing in the eight schools the Comptroller's auditors had visited during the 2008 English and Math tests. The only DOE personnel who visited those schools on those days were the audit aides who assisted the Comptroller's auditors. As in all audits, DOE expeditiously complied with the request for non-privileged documents. DOE assumed the auditors wanted to compare their observations at the eight schools they had chosen to visit with the observations of the audit



aides who had accompanied and assisted—and observed the same things as—the auditors at those eight schools. As of this point in time, the Comptroller had indicated no interest in DOE’s supplemental monitoring of schools and had asked for no monitoring checklists filed by DOE’s supplemental monitors, nor had the schedule of monitoring visits been requested.

- Months later the Comptroller asked DOE to produce the schedule of monitoring visits and the monitoring checklists filed by DOE volunteers who took part in 2008 supplemental monitoring. Again, DOE expeditiously complied with the request. As the Draft Report acknowledges, it was not until “February 2009” that the auditors for the first time “informed DOE” that they believed the audit aides were supplemental monitors, and “[t]he following day, DOE officials stated that the [aides] ‘were *not* assigned to those schools on those days [i.e., the schools and days chosen by the auditors to conduct site visits] to perform a monitoring function, and thus would not have been expected to complete a monitoring checklist’” or perform supplemental-monitor duties (p. 12).<sup>8</sup>

## VII. THE COMPTROLLER’S INACCURATE FINDINGS ON DOE’S SUPPLEMENTAL MONITORING

The auditors observed 17 test administrations during four tests: January 2008 English, March 2008 Math, April 2008 Science and January 2009 English. At no time during the four test periods did they ask to observe supplemental monitoring. After the auditors in February 2009 finally informed DOE that they were interested in supplemental monitoring, DOE invited the auditors to observe supplemental monitoring in March 2009, but the auditors declined to do so.

The Comptroller’s Draft Report concludes that DOE’s voluntary supplemental monitoring is “an effective tool that can be used to minimize the potential for improper conduct” (p. 12). But the Draft Report opines that the supplemental monitoring that takes place is “inadequate.” As is catalogued below, this opinion is based on obvious mistakes of fact and on practices DOE no longer uses and changed before 2009 tests. (DOE informed the Comptroller of the misstatements at the exit conference, but they remain in the Draft Report.)

(1) The Draft Report says that DOE does not keep track of the monitors assigned to visit schools and the submission of checklists. This statement is false for both 2008 and 2009. In 2008, DOE maintained lists of assigned visits and reviewed monitoring checklist forms submitted by monitors following their visits. And in 2009, DOE upgraded its recordkeeping.

(2) The Comptroller says DOE failed to receive monitoring checklists following monitoring visits that in fact *never occurred* because DOE *did not undertake to monitor all 269 visits that were tentatively scheduled*. Of the 128 checklists the Comptroller claims are “missing,” 86 were for schools that were on the *tentative* monitoring list, but didn’t have a monitoring visit because no

---

<sup>8</sup> Given this timeline, it is bizarre that the Draft Report (p. 12) finds it amiss that DOE turned over the audit aides’ forms to the auditors when they specifically asked for the forms, then months later turned over the 2008 supplemental monitors’ checklists when the auditors asked for them. After having assigned audit aides to assist the auditors in their visit to eight schools—and months before the auditors expressed any interest in DOE’s supplemental monitoring—DOE had no reason to doubt that the auditors’ request for forms filed by DOE central employees at only those eight audited schools pertained to what the audit aides had observed at the schools while accompanying the auditors. It wasn’t until months later that the auditors asked for the checklists filed by 2008 supplemental monitors. And it wasn’t until several months after that, in February 2009, that the auditors told DOE that they believed the audit aides were supplemental monitors (p. 12). As the Draft Report notes, DOE immediately told the auditors they were mistaken.

volunteer monitor was available to cover that school on the requisite date and time. The Comptroller's claim that monitoring checklists were filed for only 48% of schools monitored in 2008 thus is mistaken. As is noted above, of the 183 unannounced monitoring visits at elementary schools that occurred during the winter 2008 test administrations, 147 (80%) were the subject of filed monitoring checklists. And of the 111 unannounced monitoring visits to elementary schools during the most recent—winter 2009—test administrations, all 111 (100%) visits were confirmed and monitoring checklists received.

(3) The Comptroller mistakenly states that the six *audit aides* who accompanied his auditors on their eight unannounced school visits in winter 2008 were DOE *supplemental monitors*, and then criticizes their so-called monitoring because (i) monitoring checklists were filed on only six of the eight visits and (ii) the *audit aides* did not engage in the full range of activities that *supplemental monitors* are required to undertake.<sup>9</sup> But the reason the audit aides failed to engage in the activities that DOE supplemental monitors undertake is that the aides were not at the schools to monitor them. As DOE's e-mail to the Comptroller's auditors states (see Appendix G), the aides were there to "meet" and "assist you," *the auditors*. To the extent the aides' observations on those eight occasions were less comprehensive than those required of DOE's supplemental monitors, it is only because the Comptroller's audit observations at the eight schools were less exacting than those of DOE's supplemental monitors. The fact that six audit aides used monitoring checklists to report on the activities the *Comptroller's auditors observed* while the audit aides assisted them is irrelevant to the question of how many monitoring checklists were filed after DOE supplemental monitoring at schools that *were* chosen to be monitored and is also irrelevant to the actions DOE supplemental monitors undertake when they *are* monitoring schools selected for that purpose.

(4) Although some schools don't open their doors until after 7:30, the Comptroller finds it amiss that 32% of the monitoring checklists indicate that the reviewer entered the school after 7:30 a.m. The monitoring checklist lists 7:30 a.m. as the appropriate time for supplemental monitors to arrive at the school *doors*, but it doesn't require monitors to enter the school before the doors open to the public, because that can only be arranged by informing schools ahead of time that a visit is to occur, which would violate procedures in place to assure that monitoring visits are unannounced until they begin. The relevance of the time the monitor entered the school is in regard to whether the monitor viewed the removal of shrink wrap from the test booklets prior to test administration—actions that take place in most schools between 8:00 and 9:00 a.m. The monitoring checklists provide the best, direct evidence on that question. 98% of 2008 monitoring checklists reveal that the monitors were at the school in time to view removal of the shrink wrap and that the monitor did see the shrink wrap being removed.<sup>10</sup> The Comptroller found no evidence to the contrary.

(5) The Draft Audit Report dedicates almost two full pages to findings and recommendations about items on the monitoring checklist. These findings are based on the erroneous assumption—belied by the auditors' own procedures and many of the Comptroller's own findings—that investigations must in all cases be based on matters observed first-hand and cannot be based on before and after-the-fact interviews of knowledgeable individuals to determine their understanding of the rules and the procedures they followed. The Draft Report assumes that if an action cannot

---

<sup>9</sup> For example, the Draft Report states: "two of the eight monitors did not observe testing classrooms at all to ensure that the exam was administered according to State guidelines," and "some monitors did not observe classrooms while the test was in progress." (p. 12-13)

<sup>10</sup> The remaining 2% indicate the monitor was not present to verify.

be observed first-hand, filling out information about that action “renders the answers of little or no value” (p. 12). This conclusion violates all known investigative protocols.<sup>11</sup> As is true of all investigators, supplemental test monitors can’t be everywhere at once, and in some cases are prohibited by NYSED regulations from observing key activities. To assess a school’s compliance with the state-mandated testing protocols, monitors have no alternative to using a mixture of first-hand observation and interviews of knowledgeable individuals. A monitor can easily determine if an issue was sufficiently addressed at a school by asking a few simple questions. The Comptroller is correct that a few items on the 2008 list were unnecessary, but those items were removed or re-written during administration of the 2009 state tests—as the auditors would have seen, if they had accepted DOE’s invitation to observe monitoring taking place during 2009 Math tests.

### VIII. THE COMPTROLLER’S TREND AND ERASURE ANALYSIS

As is described in detail on page 2 above, the Comptroller conducted extensive statistical and forensic analyses to identify potential “inappropriate manipulation of test scores or cheating,” including (1) review of hundreds of delivery logs, letters explaining late delivery of documents and make-up exam reports; (2) forensic (erasure) analysis of hundreds of student answer documents at schools with large score increases; (3) a detailed statistical analysis of score patterns at schools in each of the city’s 32 community districts where mean scale scores increased or decreased the most; and (4) a detailed review of substantiated allegations of cheating referred to the Office of Special Investigations for the 2005-06, 2006-07, and 2007-08 school years.

After conducting these painstaking analyses, the Comptroller was “unable to draw any conclusions” of cheating, manipulation, tampering or other irregularities (p. 16). Nevertheless, the Comptroller opines that DOE’s internal controls are inadequate. The basis for this opinion is a series of misstatements of fact, some of them egregious and irresponsible:

#### Erasure analysis.

- The Comptroller’s Draft Report says that “prior to calendar year 2002, DOE performed erasure analyses” at which point “DOE stopped using” the analysis. It wasn’t the “DOE” (the Department of Education was created in 2002), but instead the former Board of Education, that started and stopped using systematic erasure analysis *prior* to 2002. The Board of Education eliminated the practice in 2001, because it was found to be a poor use of resources that didn’t turn up appreciable evidence of cheating or manipulation. The Comptroller was President of the Board of Education at the time this decision was made.
- The Comptroller’s Draft Report says DOE initially “contended” that the reason for discontinuing erasure analysis was related to NYSED’s answer key distribution policies. This too is a misrepresentation. During a discussion with audit staff, in response to questions about the history of erasure analysis, DOE officials hypothesized that NYSED’s answer key distribution policies may have been a factor, but said they weren’t certain because no one on DOE’s central test administration staff had been at DOE long enough to remember the practice. When the auditors asked DOE for more information about the prior erasure policy, DOE contacted the former head of the Division of Assessment and Accountability (DAA), who worked for the Board of Education and then DOE between 1982 and 2006 and provided the

---

<sup>11</sup> It is ironic that the Comptroller criticizes findings not based on personal observations in the midst of a lengthy critique of a procedure—supplemental monitoring—that the Comptroller and his auditors *refused* to observe in person.

information that the Comptroller requested, namely, that the systematic erasure analysis had been scaled back in 2000 and terminated in 2001 because it was unproductive.<sup>12</sup>

- Most egregious is the Comptroller’s statements about “test answer documents ... from two students from two different schools with *an excessive pattern of erasures*” (p. 14-15). The statement refers to the results of an analysis in which the audit staff used statistical methodology to select schools where 4<sup>th</sup> grade students attained the greatest gains from 2007 to 2008; randomly selected 4<sup>th</sup> grade classrooms at the schools; and reviewed the answer documents of the 336 students in those classrooms to see if there were suspicious patterns of erasures and changes of incorrect to correct answers.<sup>13</sup>
  - The statement that the two answer documents had “an excessive pattern of erasures” is odd, because a page later, the Draft Report says in no uncertain terms that the auditors’ erasure analysis did not permit “*any* conclusions” of cheating, manipulation, tampering or other evidence of irregularities (p. 16). This finding is borne out by the Comptroller’s failure to report these or any other alleged irregularities to the Office of the Special Commissioner of Investigation or to the DOE at any time during the course of the fieldwork, in repeated meetings with the DOE, in the auditors’ preliminary report or at the exit conference.
  - Also perplexing is the Comptroller use of the phrase “excessive pattern of erasures,” though the two answer sheets he refers to are admittedly from a *single* student in each of two *different* schools. As the former head of DAA told DOE, even when the former Board of Education conducted erasure analysis, a “pattern of erasures” was never deemed “excessive” or worthy of further scrutiny unless there were *multiple*—at least *five*—answer sheets with large numbers and similar patterns of answers changed from wrong to right *in the same classroom or school*. The testing expert whose opinion letter is in Appendix H uses a similar definition of “excessive erasures.” The Comptroller found no indication of a pattern of tampering or manipulation of multiple answers by anyone at any school.
  - Odder still, when the auditor staff produced their April 7, 2009 preliminary draft report in preparation for the May 4, 2009 exit conference, the report contained no reference to the two answer documents, and the auditors explicitly assured DOE’s Auditor General that there were *no cases or other evidence* of cheating or manipulation to report. By overruling his audit staff’s judgment and introducing two new cases in his final report, the Comptroller commits a blatant breach of established protocols, under which his office provides audited agencies with all factual allegations in a preliminary draft report prior to the exit conference, so the agency can investigate and respond.
  - The plot reaches its thickest point, however, when one examines the two answer sheets in question (reproduced in Appendix H). As is shown by the table below, a simple glance at

---

<sup>12</sup> The Comptroller complains that DOE didn’t produce documents on the prior erasure policy (p. 15). DOE made a diligent and expansive search for documents on the policy—which ended years ago under a prior administration—and found that none exist. It was for this reason that DOE reached out to the former head of DAA and provided the Comptroller with written confirmation of her recollections. Since 2001, DOE has reserved erasure analysis for specific cases of alleged misconduct supported by other evidence.

<sup>13</sup> Of course, most test takers check over their answers to correct mistakes, so answers changed from incorrect to correct are not by themselves suspicious. DOE asked the Comptroller to disclose the standards he used to determine what was and wasn’t suspicious, but the Comptroller declined to do so.

both answers sheets makes obvious what happened to both students, something that can happen to anyone reading questions in a test booklet and marking answers on a separate bubble sheet: The students inadvertently “got off by one or two” rows when moving answers from the question booklet to the bubble sheet, and had to erase a sequence of answers they had bubbled into the wrong rows and move them to the right rows:

- Student 1—who scored an impressive Level 4 (the highest level possible) in 2007 (the year before the test in question), and again in 2008 (the year of the test in question) and again in 2009 (the year after)—got all of the adjusted questions right after correcting the misalignment of questions and answers.
  - Student 2—who scored a high Level 2 in 2007, a low Level 3 in 2008, and a low Level 3 in 2009—did the same thing, but in his or her case got the *wrong* answer some of the time after realigning the sequence of answers. Clearly, no one supplied this student with the *right* answers, given that three pre- and post-erasure answers were the same, *wrong* answer.
- The Comptroller’s audit staff was correct in concluding prior to producing the *preliminary* draft report, which omitted the two cases, that there was no “excessive pattern of erasures” in either case. And the Comptroller was wrong in his Draft Report to violate protocol by introducing new and erroneous charges for the first time at that late stage.
  - Lest there be any doubt, DOE asked a testing expert to examine the two answer sheets in question. Her report, in Appendix H, bears out all of the conclusions above—and identifies another reason why the erasures clearly are not “excessive” or indicative of “irregularities:” Had the two students *not* made the erasures and corrections in question, each would have scored dramatically lower in 2008 than he or she did in 2007 and 2009.

Student #1	Item #	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
Erasure Mark							B	D	C	C	B	[C]	D	A	[D]	C	A	D	C	B	[C]	A	B	B	D				
Final Answer		A	C	D	D	B	<b>C</b>	<b>B</b>	B	D	C	[C]	B	C	[D]	A	D	C	A	D	[C]	B	C	A	B	B	D	A	C

Student #2	Item #	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
Erasure Mark									D	C	[C]	B	D	[A]	B	A	D	C	B	[C]	A	B	B	C	A	C			
Final Answer		A	B	C	A	B	C	B	<b>B</b>	D	[C]	C	B	<b>[A]</b>	D	A	C	B	A	D	[C]	B	C	A	B	B	B	A	C

**Yellow** indicates questions where the student initially misaligned answers bubbled onto the answer sheet by one or two rows, then erased the misaligned answers and re-bubbled the same sequence into the appropriate rows.

**Pink** indicates questions the student initially skipped and later answered after the misalignment was discovered.

**| |** indicates an answer the student did not need to erase and re-bubble because the misaligned and corrected answers were the same.

**Asterisk\*** indicates the 1 erasure out of the 31 total erasures that doesn’t fit the pattern.

**BOLD TYPE** indicates Student 2’s incorrect final answers, including three bubbled-in after the student corrected the misalignment.

**Make-up exams, late deliveries and Day Two and Day Three tests:**

- After describing his comprehensive statistical and other analyses of 2008 test score documents, logs, trends and the volume of make-up exams and late deliveries of answer documents, and after finding *no evidence* of cheating, manipulation, tampering or irregularities, the Comptroller tries to undermine the finding by stating that DOE's one-year document-retention policy kept him from examining materials for multiple years (p. 16). This is another gross misstatement. Although originally announcing that he intended to audit the 2008 state tests, the Comptroller unilaterally extended the audit to the 2009 state tests—which were the subject of nearly as many of his auditors' unannounced site observations as in 2008. The Comptroller thus had open access to a second full year of test documents, logs and data on late delivery of answer documents, make-up exams, Day Two and Day Three tests, supplemental monitoring and everything else. It was *he* who chose not to look at the second year of material and data. As for DOE's policy of storing (at great expense) the millions of pages of state testing materials generated annually for a single calendar year, the Comptroller himself found that this, and all other DOE test administration policies, are in accordance with the governing NYSED requirements and procedures (p. 1). Again, in his Draft Report, the Comptroller fails to identify any other standards, rules or practices suggesting a different retention policy.

**Other "indicators":**

- The most misleading statement in the Comptroller's Draft Report is a passage purporting to list four "indicators" that "*may* signal testing irregularities" (p. 15). The first "indicator" is the erasures on the two student answer sheets discussed above, which clearly provide no such indication. As for the second and third alleged indicators—"excessive number of make-up exams" and "late delivery of exams"—the Comptroller himself acknowledges on the next page that, after conducting meticulous analyses of both issues, he "*was unable to draw any conclusions*" that make-up exams or late deliveries were excessive or irregular in any way (p. 16). Worst of all is the fourth asserted "indicator," namely, the "significant increase in test scores" in the city. Having himself validated the integrity of city test score increases by failing, despite 18 months of exhaustive digging, to find a single "instance of cheating" or misconduct—the Comptroller's statement is a clear and irresponsible effort to demean the hard work of every child and teacher in the city in the same gratuitous and unfair way he demeans the efforts of the two 4<sup>th</sup>-grade students discussed above to bubble in what they thought were the right answers.

DOE is reviewing the Comptroller's proposed erasure, late delivery and make-up test trend analysis to determine whether it is feasible and productive, notwithstanding the consistently uninformative results such analysis produced when conducted by the former Board of Education and by the Comptroller himself in his audit. DOE is hampered in completing this review by the Comptroller's refusal to disclose the methods and standards he used and the time and resources he spent on the trend analysis he conducted.

**IX. THE COMPTROLLER'S OTHER INACCURATE AND MISLEADING FINDINGS**

The Comptroller makes a number of additional statements that are not supported or are disproved by the evidence the Comptroller developed, and makes recommendations that are not based on any discernable standards and that, in one instance, would be harmful to children:

- **Statement:** “DOE does not have a formalized process in place to make certain that substantiated cases of cheating on State tests, whether investigated by SCI or DOE’s OSI are shared with [DAAR]. Consequently, we cannot be assured [DAAR] is aware of problematic areas that need to be addressed when administering standardized tests” (p. 18).

**Response:** This statement is erroneous. DAAR is aware of allegations *during* SCI and OSI investigations. When an allegation of a testing violation is made, DAAR often is contacted first and in some cases refers the matter to SCI. Once SCI or OSI undertakes an investigation, it works closely with a DAAR specialist to investigate the allegations. The DAAR specialist has deep knowledge of testing practices, provides continuous feedback to her DAAR colleagues on potential risk areas with respect to testing misconduct and proposes changes in testing practices to improve security processes and procedures in city schools. DAAR has long used lessons learned from allegations to improve test security procedures. For example, the *Test Material Security Form*—one of DOE’s supplemental controls, which every test proctor fills out—was revised in 2008 in response to an allegation made to SCI and communicated to DAAR’s specialist. DOE also has a process under which OSI and the Administrative Trials Unit *immediately share all substantiated cases of testing improprieties with DAAR.*

- **Statements:** “[I]mproper proctoring was identified as a significant issue” because it was the case in nine substantiated cases of cheating across the three years of data the Comptroller analyzed (p. 18), DAAR should “require schools [to] assign exam proctors other than the class’s own teacher” and “periodically inform teachers and administrators of the possible consequences they can face if they are found to be involved with test tampering” (p. 19).

**Response:** *First*, the Comptroller’s assertion that 9 substantiated allegations across the 135,000 classroom administrations in which tests were separately administered and proctored during the 3 years he studied—i.e., well less than **1/100th of a percent** of the classroom administrations—is a “significant issue” is absurd on its face.<sup>14</sup> Quite the contrary is suggested. *Second*, the Comptroller’s proposal to require 7- and 8-year old children to take state high-stakes tests in rooms where no familiar adult is present is not in the best interest of children and has no basis in the governing NYSED rules. Given the young age of the students in question, the maintenance of a calm and familiar atmosphere is essential. *Third*, teachers and administrators are made fully aware of proper proctoring guidelines as a result of DOE’s supplemental control requiring each proctor to review the Test Administration Handbook and sign a statement confirming their understanding. Teachers and administrators know that if they are found to have engaged in proctoring irregularities, they face serious disciplinary action, including termination. *Finally*, and most importantly, after conducting unannounced visits at numerous schools and “observ[ing] the testing classrooms for the duration of the exam” (p. 11), the Comptroller explicitly finds that “DOE has adequate internal controls with respect to ensuring that schools are familiar with established procedures when administering New York State standardized tests at elementary schools” and that educators at the schools “complied with the State testing guidelines, DOE Handbook, and testing memoranda” (p. 1). This is hardly a justification for the precipitous and dangerous steps the Comptroller recommends.

- **Statement:** “Day Two and Three portions of the English and Math tests remain at schools for a number of days, which increases the risk of their being inappropriately manipulated” (p. 17).

---

<sup>14</sup> NYS English and Math exams are each administered to approximately 225,000 elementary students in about 11,250 separate classrooms annually: 11,250 x 4 days of testing per year (2 for English; 2 for Math) x 3 years = 135,000.

Response: The Comptroller's belief that there is a risk of improper manipulation of Day Two and Three materials is based on a single substantiated case over three years and well over a million Day Two and Day Three test booklets. That isolated instance is no basis for changing procedures that follow NYSED requirements and have worked exceedingly well to *avoid* manipulation across millions and millions of student tests. Nonetheless, DOE is considering the cost and feasibility of the Comptroller's suggestion to use "dated seals" (p. 17).

- Statement: "There is a potential for security of the exam to be compromised, since the exam is identical throughout the State and given at various times in different parts of the State" (p. 19).

Response: DOE cannot deviate from NYSED-defined assessment timeframes and goes beyond all NYSED mandated testing procedures and protocols to maintain the security of tests. NYSED determines the calendar dates and length of the testing window for all statewide assessments, and neither DOE nor other districts can unilaterally change those windows. DOE agrees that the long NYSED administration window creates test security concerns, and DOE has repeatedly urged NYSED to shorten the administration window. As a step in the recommended direction, NYSED recently conducted a statewide survey soliciting feedback on possible changes to its current test administration requirements, including timing.

In the meantime, DOE has developed its own voluntary and effective work-around to assure that NYSED testing windows don't compromise the security of DOE tests. To forestall the possibility the Comptroller hypothesizes—that students being tested late in the testing window will unfairly gain access to exam questions from students tested elsewhere in the state at an earlier point in the testing window, DOE sets the citywide administration dates at the *beginning* of the window. The Comptroller identifies no example in which DOE's local practice has failed to forestall the problem he hypothesizes, and DOE is aware of no such example.

## **X. THE COMPTROLLER'S VALID SUGGESTIONS, WHICH DOE ALREADY HAS IMPLEMENTED OR IS REVIEWING**

The Comptroller's Draft Report identifies a number of potentially valid enhancements to existing DOE processes, which DOE has either already adopted or is considering. These include:

- enhancing monitoring checklist retention and scheduling processes (implemented winter 2009)
- clarifying monitoring checklist items and clarifying monitor tasks (implemented winter 2009)
- adding a statement to the monitoring checklist instructions that mirrors statements in materials provided to schools and monitors that test-booklet shrink wrap may not be opened more than one hour before the exam (implemented winter 2009)
- formalizing process by which DAAR obtains final reports of substantiated testing allegations from OSI (OSI reports) and ATU (SCI reports) (implemented March 2009)
- asking NYSED to shorten the test administration window (implemented last several years)
- conducting cost-effective methods of erasure analysis (under consideration)
- storing returned monitoring checklists electronically (under consideration)
- analyzing longitudinal data related to late deliveries and make-up exams (under consideration)
- using "dated seals" on Day Two and Day Three materials (under consideration)

## **XI. CONCLUSION**



DOE strives for reliability in all test administration, handling, delivery, scanning, scoring and reporting and engages in 20-plus voluntary procedures, above and beyond the requirements of federal and state law, to assure test integrity, including unannounced monitoring of schools during test administration windows. The Comptroller spent 18 months and untold city resources—including on comprehensive documentary, forensic and statistical analysis—looking for evidence that test scores reported for DOE students are inaccurate, and he found *no problems* with test administration, delivery, handling, scoring, scanning and reporting in DOE schools and *no evidence* of cheating, manipulation, tampering, misconduct or other irregularities.

When the Comptroller failed to find problems with the 2008 test administration, the audit was expanded beyond its original scope to encompass 2009 tests. Again, the Comptroller found *no problems* with test administration, delivery, handling, scoring, scanning and reporting in DOE schools and *no evidence* of cheating or other misconduct.

When the Comptroller again refocused the audit, this time on the supplemental monitoring DOE voluntarily undertakes—exceeding the requirements of state law and augmenting monitoring, investigation and audits by the New York State Education Department, the Office of Special Investigations and the Comptroller himself—DOE invited the Comptroller to select schools where he could actually observe supplemental monitoring taking place. The Comptroller declined, on the ground that he did not want to conduct additional fieldwork.

And when the Comptroller's staff produced a preliminary draft report acknowledging that their erasure analysis revealed no evidence of cheating or manipulation, the Comptroller violated longstanding protocol by pulling new allegations out of the hat at the last minute when timely response was difficult—in the process subjecting two 4<sup>th</sup> grade children and their teachers and principals to charges that visual inspection of the relevant documents reveal to be false.

It is difficult to fathom the Comptroller's decision to de-emphasize his consistently favorable findings on the absence of cheating, manipulation or other irregularities in the administration of state tests and to manufacture problems where none exist. After expending a wealth of public time and resources on his audit, the Comptroller concluded that New York City educators and administrators are well-trained in all state and local requirements relating to test administration, that they administer the tests in compliance with those requirements, and that no evidence of cheating, manipulation, tampering, irregularities or misconduct was found. Yet, ignoring his own positive findings, the Comptroller focuses his Draft Report almost exclusively on minor quibbles about aspects of one of the DOE's more than 20 supplemental controls that he did not actually observe or responsibly investigate. And he raises suspicions about two 4<sup>th</sup> grade answer documents (out of nearly 2 million generated during the audit period) where it is obvious that the most mundane of bubbling-in mistakes occurred and were corrected. That the Comptroller has had to struggle this hard to find something to complain about in a comprehensive 18-month audit is a testament to the accuracy and integrity of test results showing years of impressive learning gains by New York City children and their teachers.

**Appendix A**  
**New York State ELA and Math Exam Security Processes**

	<b>NYS Requirements</b>	<b>DOE Supplemental Controls</b>
Testing Window	<ul style="list-style-type: none"> <li>▪ Schools must arrange to administer the English Language Arts and mathematics Tests to students during the specified administration week for the test.</li> <li>▪ Within each grade, schools are to administer the test on the same day to all general education students except for those students who are absent.</li> <li>▪ The makeup week, which immediately follows the administration week, is to be used for administering makeup tests to students who were absent during the primary administration week.</li> <li>▪ Scoring rubrics and answers are released at the end of the make up period.</li> </ul>	<ul style="list-style-type: none"> <li>▪ DOE designates citywide administration dates starting with the first Tuesday of the testing window (except in extenuating circumstances) to ensure that all schools citywide take the assessments on the same day. Testing at the beginning of the window and on standardized days citywide increases the security of the test.</li> <li>▪ DOE mandates that test administration must begin no later than 9:15 a.m.</li> </ul>
Dissemination of Policies and Procedures	<ul style="list-style-type: none"> <li>▪ Testing manuals outline SED requirements and are posted on NYSED website in PDF form:                             <ul style="list-style-type: none"> <li>○ <i>School Administrator's Manual</i></li> <li>○ <i>Scoring Site Operations Manual</i></li> </ul> </li> <li>▪ NYSED requests that all test proctors are provided a copy of the <i>Teacher's Directions</i> prior to the test date.</li> <li>▪ Testing memoranda outline SED requirements and are posted on NYSED website in PDF form.</li> </ul>	<ul style="list-style-type: none"> <li>▪ A Test Coordinator is designated at each school and is responsible for attending test administration meetings and disseminating information regarding test-related processes and procedures at their school.</li> <li>▪ DOE requires all testing proctors review <i>Test Administration Handbook</i> at the beginning of each school year and sign a roster to confirm understanding of all testing procedures and policies.</li> <li>▪ DOE requires that each Principal review <i>Test Administration Handbook</i> with school staff (testing proctors) and fax confirmation at the beginning of each school year.</li> </ul>

Appendix A  
New York State ELA and Math Exam Security Processes

	NYS Requirements	DOE Supplemental Controls
<p>Security of Test Materials</p>	<ul style="list-style-type: none"> <li>▪ Test materials arrive one week prior to the start of the administration window and Principals are requested to inventory materials received to ensure it matches the packing slip.</li> <li>▪ NYSED requires that test materials are maintained in shrink wrap and in a locked, secure location until the date the exam is administered.</li> <li>▪ NYSED designates Principal responsible for the security of test materials. Materials must be kept in a safe or vault and the key or combination must be maintained under strict security conditions before, during, and after the administration of the exam.</li> </ul>	<ul style="list-style-type: none"> <li>▪ Prior to each test administration, the Assessment Implementation Directors (AIDs) hold a testing coordinator's meeting to review and confirm all testing policies and procedures.</li> <li>▪ DOE prepares testing memoranda for each exam administration outlining policies and procedures and disseminates to schools and posts in PDF form online.</li> </ul>
		<p><u>School-Based Procedures:</u></p> <ul style="list-style-type: none"> <li>▪ At each school, the <i>Test Material Security Form</i> is used to track distribution and return of exams on testing day; the number of tests being picked-up or dropped off are counted, signed for and initialed by both the teacher and the testing coordinator.</li> <li>▪ DOE requires supervisors/test coordinators may not open the shrink-wrapped tests to make class sets until 1 hour before the administration. Teachers/proctors are given 15 minutes before testing begins to open the sealed packages and distribute them.</li> </ul> <p><u>Other Procedures:</u></p> <ul style="list-style-type: none"> <li>▪ Day 1 answer documents (Multiple Choice bubble sheets) are collected by 3:00 p.m. on the day of the test administration and delivered to a secure, centralized location to ensure test security.</li> <li>▪ Day 1 Make-Up exams are delivered to the centralized secure location on the last day of the test administration.</li> </ul>

Appendix A  
New York State ELA and Math Exam Security Processes

	NYS Requirements	DOE Supplemental Controls
Monitoring of Test Sites	<ul style="list-style-type: none"> <li>▪ NYSED conducts a small number of unannounced site visits to schools across the state to ensure test materials are securely maintained.</li> <li>▪ NYSED provides security guidelines in the <i>School Administrator's Manual</i> and in the <i>Scoring Site Operations Manual</i> to facilitate secure handling of test materials.</li> </ul>	<ul style="list-style-type: none"> <li>▪ DOE coordinates unannounced test monitoring visits at randomly selected schools during the NYS ELA and math test administrations. Uniform monitoring criteria are applied through use of a standardized monitoring checklist form that monitors complete and submit following school visit.</li> <li>▪ Log sheets are maintained at each ISC to document return of answer documents by schools and are faxed to scan center while documents are being transported via secure courier to scan center.</li> <li>▪ DOE requires Principals to submit a letter if test materials are returned late.</li> <li>▪ Log sheets are maintained at the scan center to track secure test materials received and are checked against the ISC log sheets to ensure materials have been securely transferred.</li> </ul>
Documentation of Test Material Handoffs		
Scoring of Constructed Response Questions	<ul style="list-style-type: none"> <li>▪ NYSED posts <i>Scoring Site Operations Manual</i> and distributes <i>Scoring Leader Handbook</i> which outline scoring requirements, including: <ul style="list-style-type: none"> <li>○ Scorers must be NYS-certified educators</li> <li>○ Scorers must be trained on NYSED-provided training materials and rubrics and pass CAS prior to scoring</li> <li>○ Each scoring team must have 3 scorers for each test</li> </ul> </li> <li>▪ NYSED allows for school-based or regional/centralized scoring.</li> </ul>	<ul style="list-style-type: none"> <li>▪ DOE utilizes a centralized scoring model in which teachers from schools citywide score randomized papers.</li> <li>▪ Scorers are not allowed to score their own school's tests.</li> <li>▪ Each test must be scored by three raters with the items divided among scoring teams.</li> <li>▪ Materials are maintained in secure, locked centrally-managed facility.</li> <li>▪ In addition to the state-mandated scoring site procedures, DOE also utilizes anchor papers to monitor scoring quality.</li> </ul>

Appendix A  
New York State ELA and Math Exam Security Processes

	NYS Requirements	DOE Supplemental Controls
Testing Irregularities	<ul style="list-style-type: none"> <li>▪ Materials must be kept confidential and in a secure location.</li> <li>▪ NYSED requires that any security breaches or testing irregularities be documented and sent to Office of State Assessment within 24 hours.</li> </ul>	<ul style="list-style-type: none"> <li>▪ During test administration period, if test irregularities occur at the school, such as inappropriate proctoring, or tampering with secure materials, Principal is immediately notified and incident is reported to SCI, ISC, and NYSED. All testing documentation (test memoranda, <i>Administration Handbook</i>) clearly outline these escalation steps.</li> <li>▪ During scoring, irregularities are brought to the attention of the table leader who notifies site supervisor (such as handwriting discrepancies, excessive erasures). If a determination is made that further action is warranted, SCI, ISC, DAAR, and SED are notified immediately and investigation may be pursued by OSI or by SCI.</li> </ul>

**Appendix B – Auditor School Visits  
NYS ELA and Math\*  
2008 & 2009 Exam Administrations**

<b>Test</b>	<b>School</b>	<b>Borough</b>	<b>Date of Observation</b>
ELA	PS230	Brooklyn	1/8/08
Math	PS112	Queens	3/4/08
Math	PS130	Manhattan	3/5/08
Math	PS121	Bronx	3/5/08
Math	PS179	Brooklyn	3/5/08
Math	PS082	Queens	3/6/08
Math	PS097	Bronx	3/6/08
Math	PS234	Manhattan	3/6/08
ELA	PS111	Queens	1/13/09
ELA	PS209	Brooklyn	1/13/09
ELA	PS001	Manhattan	1/14/09
ELA	PS048	Manhattan	1/14/09
ELA	PS253	Brooklyn	1/14/09
ELA	PS117	Queens	1/15/09
ELA	PS078	Bronx	1/15/09

\*The auditors' school visits included PS 89 in the Bronx and PS 15 in Manhattan to observe NYS Science exam administration on April 30, 2008. The Draft Report does not claim that test monitoring was observed during these observations, therefore the two Science observations are excluded from the above list.

Appendix C

THE NEW YORK CITY DEPARTMENT OF EDUCATION  
TEST ADMINISTRATION  
SECURITY MONITORING FORM

MATHEMATICS-2009

- Monitors must arrive at assigned school by 7:30 a.m. and stay until testing is completed.
- Tests will be in shrink-wrapped packages.
- Monitors should examine shrink-wrap on the morning of the test.
- Building Supervisors/Testing Coordinators may open shrink-wrap packages 60 minutes before test administration.
- Proctors may open class sets/shrink-wrapped packages in the classroom 15 minutes before the test begins.
- All tests will be administered under strict time limits. Monitors should ensure the enforcement of time limits. Some students will be entitled to extended time based on their IEPs or 504 Plans.
- Monitors will use the attached security monitoring form during the site visit.
- Report any irregularities to Grace Pepe at (718) 349-5627.
- All security monitoring forms must be delivered to Ardora Holden at TWEED Room #309 by close of business on the day of the test or to the Assessment Implementation Director (AID) at the Integrated Service Center (ISC).
- Please **actively monitor** all aspects of the entire test administration according to the security monitoring form.

Thank you

**THE NEW YORK CITY DEPARTMENT OF EDUCATION  
TEST ADMINISTRATION  
SECURITY MONITORING FORM  
MATHEMATICS-2009**

**District:** \_\_\_\_\_ **School:** \_\_\_\_\_  
**Date:** \_\_\_\_\_ **Time Arrived:** \_\_\_\_\_  
**Test:** \_\_\_\_\_ **Time Departed:** \_\_\_\_\_

**Name of Reviewer:** \_\_\_\_\_

**Signature of Reviewer:** \_\_\_\_\_

For every "No" response, please provide a comment/explanation that clearly identifies the location (room) and the name of person (principal, teacher, student) involved.

**If a "No" response is entered for any item, this report must be faxed immediately to Grace Pepe at (718) 349-5723 or (718) 349-5642.**

**All reports must be delivered to Ardora Holden at TWEED Room 309 by close of business on the day of the test or to the Assessment Implementation Director (AID) at the Integrated Service Center (ISC).**

**If there are any questions during the review, please contact Grace Pepe at (718) 349-5627.**



**THE NEW YORK CITY DEPARTMENT OF EDUCATION  
TEST ADMINISTRATION  
SECURITY MONITORING FORM**

**MATHEMATICS-2009**

**PRE-ADMINISTRATION REVIEW**

**District:** \_\_\_\_\_ **School:** \_\_\_\_\_ **Date:** \_\_\_\_\_ **Test:** \_\_\_\_\_

ISSUE	YES/NO/ NA	COMMENT/EXPLANATION
The test booklets are stored in secured location as appropriate.		
Students and parents were informed about the test a few days in advance.		
Schools have assigned appropriate staff, such as hall proctors.		
Rooms have been set aside for latecomers and students who are not to be tested.		
Teachers are familiar with regulations concerning the exam. (i.e., timing, directions to be read, and special education issues).		
Shrink-wrapped packets of Test Booklets are not opened until the day of test.		

THE NEW YORK CITY DEPARTMENT OF EDUCATION  
TEST ADMINISTRATION  
SECURITY MONITORING FORM

MATHEMATICS-2009

TEST ADMINISTRATION CHECKLIST

District: \_\_\_\_\_ School: \_\_\_\_\_ Date: \_\_\_\_\_ Test: \_\_\_\_\_

ISSUE	YES/NO/ NA	COMMENT/EXPLANATION
On the day of the test, the test booklets/class sets are kept in a secure area inaccessible to unauthorized persons until they are distributed to teachers and proctors.		
The testing rooms are adequately lighted and ventilated and free from noise and other distractions.		
The teacher records the starting time on a chalkboard visible to all pupils in the room. There is also a clock visible to the students, or the teacher periodically updates the time remaining on the chalkboard.		
The teacher has: <ul style="list-style-type: none"> <li>• a copy of the test booklet for demonstration purposes; and</li> <li>• a supply of all other materials to be distributed to students.</li> </ul>		

THE NEW YORK CITY DEPARTMENT OF EDUCATION  
TEST ADMINISTRATION  
SECURITY MONITORING FORM

MATHEMATICS-2009

TEST ADMINISTRATION CHECKLIST

District: \_\_\_\_\_ School: \_\_\_\_\_ Date: \_\_\_\_\_ Test: \_\_\_\_\_

ISSUE	YES/NO/ NA	COMMENT/EXPLANATION
Desks are cleared of all books, paper, etc.		
All materials that might help the students have been covered or erased.		
Students are seated so that they cannot see other students' answer sheets.		
The teacher distributes appropriate materials to each student: <ul style="list-style-type: none"> <li>• test booklet; and</li> <li>• answer documents</li> <li>• manipulatives.</li> </ul>		
Teachers are actively proctoring.		

THE NEW YORK CITY DEPARTMENT OF EDUCATION  
TEST ADMINISTRATION  
SECURITY MONITORING FORM

MATHEMATICS-2009

TEST ADMINISTRATION CHECKLIST

District: \_\_\_\_\_ School: \_\_\_\_\_ Date: \_\_\_\_\_ Test: \_\_\_\_\_

ISSUE	YES/NO/ NA	COMMENT/EXPLANATION
The teacher monitors the test administration ensuring that all students are working independently.		
The teacher maintains a quiet testing environment with a minimum of distractions. Students that have completed the test are required to sit quietly at their desks or to exit the testing room so as not to disturb others.		
The teacher makes sure that each student has turned in a test booklet and answer sheet, then follows the procedures established to ensure that answer papers are forwarded to the test coordinator for proper packaging.		

THE NEW YORK CITY DEPARTMENT OF EDUCATION  
TEST ADMINISTRATION  
SECURITY MONITORING FORM

MATHEMATICS-2009

TEST ADMINISTRATION CHECKLIST

District: \_\_\_\_\_ School: \_\_\_\_\_ Date: \_\_\_\_\_ Test: \_\_\_\_\_

ISSUE	YES/NO/ NA	COMMENT/EXPLANATION
Teachers sign a receipt/roster when they pick up and return their exams.		
All classroom door windows are uncovered.		
The principal has signed the receipt/roster with teacher signatures of pick-up and return of class set of test booklets.		
All exam booklets that have been distributed have been accounted for and put in a secure locked location after the test is completed.		

Appendix D - List of Completed ELA and Math Monitoring Visits  
Grades 3-5, 2008

School	Exam Administration Monitored	Date of Test Administration Monitored
01M034	ELA	1/8/2008
02M011	ELA	1/8/2008
02M040	ELA	1/8/2008
02M089	ELA	1/8/2008
02M124	ELA	1/8/2008
02M347	ELA	1/8/2008
03M076	ELA	1/8/2008
03M087	ELA	1/8/2008
03M334	ELA	1/8/2008
05M161	ELA	1/8/2008
06M005	ELA	1/8/2008
06M173	ELA	1/8/2008
07X005	ELA	1/8/2008
07X01	ELA	1/8/2008
07X203	ELA	1/8/2008
09X004	ELA	1/8/2008
10X008	ELA	1/8/2008
10X024	ELA	1/8/2008
11X078	ELA	1/8/2008
15K058	ELA	1/8/2008
15K124	ELA	1/8/2008
15K154	ELA	1/8/2008
18K268	ELA	1/8/2008
19K072	ELA	1/8/2008
19K345	ELA	1/8/2008
20K185	ELA	1/8/2008
23K073	ELA	1/8/2008
24Q007	ELA	1/8/2008
24Q019	ELA	1/8/2008
25Q022	ELA	1/8/2008
26Q115	ELA	1/8/2008
26Q188	ELA	1/8/2008
27Q104	ELA	1/8/2008
27Q253	ELA	1/8/2008
28Q030	ELA	1/8/2008
28Q040	ELA	1/8/2008
28Q268	ELA	1/8/2008
29Q118	ELA	1/8/2008
30Q017	ELA	1/8/2008
31R023	ELA	1/8/2008
31R036	ELA	1/8/2008
31R046	ELA	1/8/2008

Appendix D - List of Completed ELA and Math Monitoring Visits  
Grades 3-5, 2008

School	Exam Administration Monitored	Date of Test Administration Monitored
31R055	ELA	1/8/2008
31R080	ELA	1/8/2008
01M110	ELA	1/9/2008
02M006	ELA	1/9/2008
02M089	ELA	1/9/2008
02M150	ELA	1/9/2008
03M075	ELA	1/9/2008
03M076	ELA	1/9/2008
03M333	ELA	1/9/2008
03M334	ELA	1/9/2008
04M206	ELA	1/9/2008
05M129	ELA	1/9/2008
06M005	ELA	1/9/2008
06M187	ELA	1/9/2008
07X156	ELA	1/9/2008
08X048	ELA	1/9/2008
08X069	ELA	1/9/2008
08X182	ELA	1/9/2008
09X090	ELA	1/9/2008
10X095	ELA	1/9/2008
10X205	ELA	1/9/2008
10X246	ELA	1/9/2008
11X068	ELA	1/9/2008
11X083	ELA	1/9/2008
11X111	ELA	1/9/2008
13K009	ELA	1/9/2008
15K015	ELA	1/9/2008
15K124	ELA	1/9/2008
15K321	ELA	1/9/2008
17K241	ELA	1/9/2008
18K135	ELA	1/9/2008
19K072	ELA	1/9/2008
20K104	ELA	1/9/2008
23K150	ELA	1/9/2008
24Q089	ELA	1/9/2008
26Q115	ELA	1/9/2008
26Q191	ELA	1/9/2008
27Q223	ELA	1/9/2008
28Q268	ELA	1/9/2008
29Q037	ELA	1/9/2008
30Q076	ELA	1/9/2008
30Q148	ELA	1/9/2008
30Q149	ELA	1/9/2008
31R031	ELA	1/9/2008

Appendix D - List of Completed ELA and Math Monitoring Visits  
Grades 3-5, 2008

School	Exam Administration Monitored	Date of Test Administration Monitored
31R046	ELA	1/9/2008
31R080	ELA	1/9/2008
02M040	ELA	1/10/2008
02M124	ELA	1/10/2008
02M183	ELA	1/10/2008
03M009	ELA	1/10/2008
03M333	ELA	1/10/2008
03M859	ELA	1/10/2008
06M098	ELA	1/10/2008
06M173	ELA	1/10/2008
08X036	ELA	1/10/2008
09X028	ELA	1/10/2008
10X008	ELA	1/10/2008
10X020	ELA	1/10/2008
10X310	ELA	1/10/2008
10X310	ELA	1/10/2008
12X134	ELA	1/10/2008
15k010	ELA	1/10/2008
15K058	ELA	1/10/2008
15K124	ELA	1/10/2008
16K028	ELA	1/10/2008
19K007	ELA	1/10/2008
19K072	ELA	1/10/2008
22K217	ELA	1/10/2008
22K222	ELA	1/10/2008
24Q089	ELA	1/10/2008
25Q154	ELA	1/10/2008
28Q268	ELA	1/10/2008
29Q116	ELA	1/10/2008
30Q017	ELA	1/10/2008
31R046	ELA	1/10/2008
31R053	ELA	1/10/2008
32K123	ELA	1/10/2008
02M006	Math	3/4/2008
02M040	Math	3/4/2008
02M116	Math	3/4/2008
03M087	Math	3/4/2008
03M334	Math	3/4/2008
06M187	Math	3/4/2008
07X157	Math	3/4/2008
07X161	Math	3/4/2008
08X140	Math	3/4/2008
10X020	Math	3/4/2008
10X081	Math	3/4/2008



Appendix D - List of Completed ELA and Math Monitoring Visits  
Grades 3-5, 2008

School	Exam Administration Monitored	Date of Test Administration Monitored
10X095	Math	3/4/2008
10X226	Math	3/4/2008
12X006	Math	3/4/2008
15K124	Math	3/4/2008
15K154	Math	3/4/2008
16K005	Math	3/4/2008
19K065	Math	3/4/2008
24Q016	Math	3/4/2008
26Q205	Math	3/4/2008
27Q114	Math	3/4/2008
28Q268	Math	3/4/2008
29Q131	Math	3/4/2008
29Q270	Math	3/4/2008
31R046	Math	3/4/2008
02M006	Math	3/5/2008
02M124	Math	3/5/2008
03M334	Math	3/5/2008
06M173	Math	3/5/2008
09X035	Math	3/5/2008
09X042	Math	3/5/2008
10X008	Math	3/5/2008
11X076	Math	3/5/2008
15K124	Math	3/5/2008
15K321	Math	3/5/2008
16K335	Math	3/5/2008
17K161	Math	3/5/2008
18K276	Math	3/5/2008
20K102	Math	3/5/2008
27Q114	Math	3/5/2008
28Q268	Math	3/5/2008
30Q017	Math	3/5/2008
30Q084	Math	3/5/2008
31R046	Math	3/5/2008
02M006	Math	3/6/2008
02M040	Math	3/6/2008
02M150	Math	3/6/2008
06M048	Math	3/6/2008
06M187	Math	3/6/2008
10X095	Math	3/6/2008
10X340	Math	3/6/2008
11X078	Math	3/6/2008
12X050	Math	3/6/2008
12X067	Math	3/6/2008
15k010	Math	3/6/2008

Appendix D - List of Completed ELA and Math Monitoring Visits  
Grades 3-5, 2008

School	Exam Administration Monitored	Date of Test Administration Monitored
15K124	Math	3/6/2008
17K221	Math	3/6/2008
22K139	Math	3/6/2008
27Q114	Math	3/6/2008
27Q114	Math	3/6/2008
28Q030	Math	3/6/2008
29Q138	Math	3/6/2008
31R046	Math	3/6/2008
32K086	Math	3/6/2008

Appendix E - Completed ELA and Math Monitoring Visits  
Grades 3-5, 2009

SCHOOL	TEST	DATE
01M015	ELA 3,4,5	01.13.09
02M124	ELA 3,4,5	01.13.09
06M048	ELA 3,4,5	01.13.09
09X004	ELA 3,4,5	01.13.09
09X230	ELA 3,4,5	01.13.09
10X032	ELA 3,4,5	01.13.09
10X081	ELA 3,4,5	01.13.09
11X105	ELA 3,4,5	01.13.09
12X044	ELA 3,4,5	01.13.09
25Q120	ELA 3,4,5	01.13.09
27Q064	ELA 3,4,5	01.13.09
30Q017	ELA 3,4,5	01.13.09
30Q069	ELA 3,4,5	01.13.09
30Q171	ELA 3,4,5	01.13.09
31R055	ELA 3,4,5	01.13.09
75K140	ELA 3,4,5	01.13.09
75K368	ELA 3,4,5	01.13.09
75K369	ELA 3,4,5	01.13.09
75K369	ELA 3,4,5	01.13.09
75M035	ELA 3,4,5	01.13.09
75M094	ELA 3,4,5	01.13.09
75Q004	ELA 3,4,5	01.13.09
75Q009	ELA 3,4,5	01.13.09
75Q075	ELA 3,4,5	01.13.09
75Q256	ELA 3,4,5	01.13.09
75R025	ELA 3,4,5	01.13.09
75X012	ELA 3,4,5	01.13.09
75X186	ELA 3,4,5	01.13.09
75X352	ELA 3,4,5	01.13.09
75X723	ELA 3,4,5	01.13.09
01M020	ELA 3,4,5	01.14.09
02M124	ELA 3,4,5	01.14.09
02M225	ELA 3,4,5	01.14.09
08X036	ELA 3,4,5	01.14.09
09X011	ELA 3,4,5	01.14.09
09X230	ELA 3,4,5	01.14.09
10X024	ELA 3,4,5	01.14.09
11x153	ELA 3,4,5	01.14.09
18K268	ELA 3,4,5	01.14.09
27Q124	ELA 3,4,5	01.14.09

Appendix E - Completed ELA and Math Monitoring Visits  
Grades 3-5, 2009

SCHOOL	TEST	DATE
28Q139	ELA 3,4,5	01.14.09
30Q234	ELA 3,4,5	01.14.09
31R042	ELA 3,4,5	01.14.09
75K036	ELA 3,4,5	01.14.09
75K081	ELA 3,4,5	01.14.09
75K140	ELA 3,4,5	01.14.09
75K231	ELA 3,4,5	01.14.09
75K368	ELA 3,4,5	01.14.09
75K369	ELA 3,4,5	01.14.09
75K771	ELA 3,4,5	01.14.09
75M169	ELA 3,4,5	01.14.09
75M361	ELA 3,4,5	01.14.09
75M811	ELA 3,4,5	01.14.09
75Q004	ELA 3,4,5	01.14.09
75Q025	ELA 3,4,5	01.14.09
75Q075	ELA 3,4,5	01.14.09
75Q224	ELA 3,4,5	01.14.09
75R373	ELA 3,4,5	01.14.09
75X010	ELA 3,4,5	01.14.09
75X016	ELA 3,4,5	01.14.09
75X168	ELA 3,4,5	01.14.09
01M315	ELA 3,4,5	01.15.09
02M225	ELA 3,4,5	01.15.09
09X073	ELA 3,4,5	01.15.09
09X230	ELA 3,4,5	01.15.09
13K054	ELA 3,4,5	01.15.09
27Q045	ELA 3,4,5	01.15.09
31R056	ELA 3,4,5	01.15.09
02M124	MATH 3,4,5	03.04.09
02M198	MATH 3,4,5	03.04.09
07X001	MATH 3,4,5	03.04.09
07X031	MATH 3,4,5	03.04.09
08X119	MATH 3,4,5	03.04.09
10X081	MATH 3,4,5	03.04.09
11X068	MATH 3,4,5	03.04.09
11X189	MATH 3,4,5	03.04.09
25Q164	MATH 3,4,5	03.04.09
30Q112	MATH 3,4,5	03.04.09
31R004	MATH 3,4,5	03.04.09

Appendix E - Completed ELA and Math Monitoring Visits  
Grades 3-5, 2009

SCHOOL	TEST	DATE
31R022	MATH 3,4,5	03.04.09
75K004	MATH 3,4,5	03.04.09
75K771	MATH 3,4,5	03.04.09
75M811	MATH 3,4,5	03.04.09
75Q075	MATH 3,4,5	03.04.09
75X010	MATH 3,4,5	03.04.09
01M315	MATH 3,4,5	03.05.09
02M124	MATH 3,4,5	03.05.09
02M198	MATH 3,4,5	03.05.09
08X069	MATH 3,4,5	03.05.09
08X119	MATH 3,4,5	03.05.09
09X022	MATH 3,4,5	03.05.09
11X041	MATH 3,4,5	03.05.09
75K036	MATH 3,4,5	03.05.09
75K140	MATH 3,4,5	03.05.09
75K192	MATH 3,4,5	03.05.09
75K231	MATH 3,4,5	03.05.09
75K369	MATH 3,4,5	03.05.09
75M094	MATH 3,4,5	03.05.09
75M094	MATH 3,4,5	03.05.09
75M169	MATH 3,4,5	03.05.09
75M169	MATH 3,4,5	03.05.09
75M811	MATH 3,4,5	03.05.09
75Q023	MATH 3,4,5	03.05.09
75Q224	MATH 3,4,5	03.05.09
75X352	MATH 3,4,5	03.05.09
75X352	MATH 3,4,5	03.05.09
75X723	MATH 3,4,5	03.05.09
75X723	MATH 3,4,5	03.05.09
02M124	MATH 3,4,5	03.06.09
09X110	MATH 3,4,5	03.06.09
11X153	MATH 3,4,5	03.06.09

Appendix F  
Preliminary Schedule of ELA and Math Monitoring Visits  
Grades 3-5, 2008

School	Exam Administration	Date of Tentatively Scheduled Monitoring Visit
01M015	ELA	1/8/2008
01M019	ELA	1/8/2008
01M034	ELA	1/8/2008
02M011	ELA	1/8/2008
02M040	ELA	1/8/2008
02M059	ELA	1/8/2008
02M089	ELA	1/8/2008
02M124	ELA	1/8/2008
02M234	ELA	1/8/2008
02M347	ELA	1/8/2008
03M076	ELA	1/8/2008
03M087	ELA	1/8/2008
03M334	ELA	1/8/2008
05M161	ELA	1/8/2008
06M005	ELA	1/8/2008
06M173	ELA	1/8/2008
07X005	ELA	1/8/2008
07X01	ELA	1/8/2008
07X203	ELA	1/8/2008
08X119	ELA	1/8/2008
09X004	ELA	1/8/2008
10X008	ELA	1/8/2008
10X024	ELA	1/8/2008
10X246	ELA	1/8/2008
11X016	ELA	1/8/2008
11X078	ELA	1/8/2008
15K058	ELA	1/8/2008
15K124	ELA	1/8/2008
15K154	ELA	1/8/2008
15K154	ELA	1/8/2008
18K268	ELA	1/8/2008
19K072	ELA	1/8/2008
19K345	ELA	1/8/2008
20K185	ELA	1/8/2008
22K251	ELA	1/8/2008
23K073	ELA	1/8/2008
24Q007	ELA	1/8/2008
24Q019	ELA	1/8/2008
25Q022	ELA	1/8/2008
26Q115	ELA	1/8/2008
26Q188	ELA	1/8/2008
27Q104	ELA	1/8/2008

Appendix F  
Preliminary Schedule of ELA and Math Monitoring Visits  
Grades 3-5, 2008

School	Exam Administration	Date of Tentatively Scheduled Monitoring Visit
27Q253	ELA	1/8/2008
28Q030	ELA	1/8/2008
28Q040	ELA	1/8/2008
28Q268	ELA	1/8/2008
29Q118	ELA	1/8/2008
30Q017	ELA	1/8/2008
31R023	ELA	1/8/2008
31R036	ELA	1/8/2008
31R042	ELA	1/8/2008
31R046	ELA	1/8/2008
31R055	ELA	1/8/2008
31R080	ELA	1/8/2008
01M019	ELA	1/9/2008
01M063	ELA	1/9/2008
01M110	ELA	1/9/2008
02M006	ELA	1/9/2008
02M089	ELA	1/9/2008
02M116	ELA	1/9/2008
02M150	ELA	1/9/2008
03M075	ELA	1/9/2008
03M076	ELA	1/9/2008
03M333	ELA	1/9/2008
03M334	ELA	1/9/2008
04M206	ELA	1/9/2008
05M129	ELA	1/9/2008
06M005	ELA	1/9/2008
06M187	ELA	1/9/2008
07X156	ELA	1/9/2008
08X036	ELA	1/9/2008
08X048	ELA	1/9/2008
08X069	ELA	1/9/2008
08X182	ELA	1/9/2008
09X090	ELA	1/9/2008
10X095	ELA	1/9/2008
10X205	ELA	1/9/2008
10X246	ELA	1/9/2008
11X068	ELA	1/9/2008
11X083	ELA	1/9/2008
11X096	ELA	1/9/2008
11X111	ELA	1/9/2008
13K009	ELA	1/9/2008
15K015	ELA	1/9/2008
15K124	ELA	1/9/2008
15K321	ELA	1/9/2008

Appendix F  
Preliminary Schedule of ELA and Math Monitoring Visits  
Grades 3-5, 2008

School	Exam Administration	Date of Tentatively Scheduled Monitoring Visit
17K241	ELA	1/9/2008
18K135	ELA	1/9/2008
19K072	ELA	1/9/2008
20K104	ELA	1/9/2008
22K251	ELA	1/9/2008
23K150	ELA	1/9/2008
24Q089	ELA	1/9/2008
25Q032	ELA	1/9/2008
26Q115	ELA	1/9/2008
26Q191	ELA	1/9/2008
27Q223	ELA	1/9/2008
28Q268	ELA	1/9/2008
29Q037	ELA	1/9/2008
30Q076	ELA	1/9/2008
30Q148	ELA	1/9/2008
30Q149	ELA	1/9/2008
31R031	ELA	1/9/2008
31R046	ELA	1/9/2008
31R055	ELA	1/9/2008
31R080	ELA	1/9/2008
01M019	ELA	1/10/2009
01M020	ELA	1/10/2009
02M001	ELA	1/10/2009
02M040	ELA	1/10/2009
02M059	ELA	1/10/2009
02M089	ELA	1/10/2009
02M124	ELA	1/10/2009
02M183	ELA	1/10/2009
03M009	ELA	1/10/2009
03M333	ELA	1/10/2009
03M859	ELA	1/10/2009
06M098	ELA	1/10/2009
06M173	ELA	1/10/2009
08X036	ELA	1/10/2009
08X100	ELA	1/10/2009
09X028	ELA	1/10/2009
09X126	ELA	1/10/2009
10X008	ELA	1/10/2009
10X020	ELA	1/10/2009
10X059	ELA	1/10/2009
10X310	ELA	1/10/2009
10X310	ELA	1/10/2009
12X134	ELA	1/10/2009
15k010	ELA	1/10/2009



Appendix F  
Preliminary Schedule of ELA and Math Monitoring Visits  
Grades 3-5, 2008

School	Exam Administration	Date of Tentatively Scheduled Monitoring Visit
15K058	ELA	1/10/2009
15K124	ELA	1/10/2009
16K028	ELA	1/10/2009
19K007	ELA	1/10/2009
19K072	ELA	1/10/2009
22K217	ELA	1/10/2009
22K222	ELA	1/10/2009
22K251	ELA	1/10/2009
23K298	ELA	1/10/2009
24Q089	ELA	1/10/2009
25Q154	ELA	1/10/2009
28Q268	ELA	1/10/2009
29Q116	ELA	1/10/2009
30Q017	ELA	1/10/2009
31R004	ELA	1/10/2009
31R046	ELA	1/10/2009
31R053	ELA	1/10/2009
32K123	ELA	1/10/2009
01M019	Math	3/4/2008
02M003	Math	3/4/2008
02M006	Math	3/4/2008
02M040	Math	3/4/2008
02M041	Math	3/4/2008
02M059	Math	3/4/2008
02M116	Math	3/4/2008
02M137	Math	3/4/2008
03M084	Math	3/4/2008
03M087	Math	3/4/2008
03M334	Math	3/4/2008
06M187	Math	3/4/2008
07X049	Math	3/4/2008
07X157	Math	3/4/2008
07X161	Math	3/4/2008
08X140	Math	3/4/2008
08X146	Math	3/4/2008
10X020	Math	3/4/2008
10X081	Math	3/4/2008
10X095	Math	3/4/2008
10X226	Math	3/4/2008
11X083	Math	3/4/2008
12X006	Math	3/4/2008
12X047	Math	3/4/2008
15K015	Math	3/4/2008
15K124	Math	3/4/2008

Appendix F  
Preliminary Schedule of ELA and Math Monitoring Visits  
Grades 3-5, 2008

School	Exam Administration	Date of Tentatively Scheduled Monitoring Visit
15K154	Math	3/4/2008
16K005	Math	3/4/2008
18K066	Math	3/4/2008
19K013	Math	3/4/2008
19K065	Math	3/4/2008
19K159	Math	3/4/2008
20K069	Math	3/4/2008
24Q016	Math	3/4/2008
26Q115	Math	3/4/2008
26Q205	Math	3/4/2008
27Q114	Math	3/4/2008
27Q183	Math	3/4/2008
28Q268	Math	3/4/2008
29Q131	Math	3/4/2008
29Q270	Math	3/4/2008
30Q166	Math	3/4/2008
31R006	Math	3/4/2008
31R041	Math	3/4/2008
31R046	Math	3/4/2008
01M019	Math	3/5/2008
02M006	Math	3/5/2008
02M042	Math	3/5/2008
02M059	Math	3/5/2008
02M124	Math	3/5/2008
02M212	Math	3/5/2008
03M087	Math	3/5/2008
03M333	Math	3/5/2008
03M334	Math	3/5/2008
06M173	Math	3/5/2008
09X035	Math	3/5/2008
09X042	Math	3/5/2008
09X070	Math	3/5/2008
10X008	Math	3/5/2008
11X016	Math	3/5/2008
11X076	Math	3/5/2008
11X106	Math	3/5/2008
11X153	Math	3/5/2008
15K058	Math	3/5/2008
15K124	Math	3/5/2008
15K321	Math	3/5/2008
16K335	Math	3/5/2008
17K161	Math	3/5/2008
18K066	Math	3/5/2008
18K276	Math	3/5/2008

Appendix F  
Preliminary Schedule of ELA and Math Monitoring Visits  
Grades 3-5, 2008

School	Exam Administration	Date of Tentatively Scheduled Monitoring Visit
19K013	Math	3/5/2008
19K260	Math	3/5/2008
20K102	Math	3/5/2008
25Q029	Math	3/5/2008
25Q201	Math	3/5/2008
27Q114	Math	3/5/2008
28Q268	Math	3/5/2008
29Q038	Math	3/5/2008
30Q017	Math	3/5/2008
30Q084	Math	3/5/2008
30Q171	Math	3/5/2008
31R046	Math	3/5/2008
31R069	Math	3/5/2008
31R080	Math	3/5/2008
01M019	Math	3/6/2008
02M006	Math	3/6/2008
02M040	Math	3/6/2008
02M051	Math	3/6/2008
02M150	Math	3/6/2008
02M225	Math	3/6/2008
03M166	Math	3/6/2008
03M333	Math	3/6/2008
06M048	Math	3/6/2008
06M187	Math	3/6/2008
07X065	Math	3/6/2008
08X119	Math	3/6/2008
10X095	Math	3/6/2008
10X340	Math	3/6/2008
11X078	Math	3/6/2008
12X050	Math	3/6/2008
12X067	Math	3/6/2008
15k010	Math	3/6/2008
15K010	Math	3/6/2008
15K015	Math	3/6/2008
15K124	Math	3/6/2008
17K221	Math	3/6/2008
18K066	Math	3/6/2008
19K013	Math	3/6/2008
19K345	Math	3/6/2008
22K139	Math	3/6/2008
27Q114	Math	3/6/2008
27Q114	Math	3/6/2008
28Q030	Math	3/6/2008
28Q268	Math	3/6/2008

Appendix F  
Preliminary Schedule of ELA and Math Monitoring Visits  
Grades 3-5, 2008

School	Exam Administration	Date of Tentatively Scheduled Monitoring Visit
29Q138	Math	3/6/2008
30Q234	Math	3/6/2008
31R003	Math	3/6/2008
31R041	Math	3/6/2008
31R046	Math	3/6/2008
31R080	Math	3/6/2008
32K086	Math	3/6/2008

**Appendix G – School Observation Request by NYC Comptroller's Office**

-----Original Message-----

From: [REDACTED]  
Sent: Thursday, February 28, 2008 12:57 PM  
To: [REDACTED]  
Cc: [REDACTED]  
(49M052); [REDACTED]  
Subject: RE: Audit of DOE's Administration of NYS Standardized Exams

The following staff will meet you at the principal's office at the schools at 7:30am to assist you with the monitoring:

March 4th	30Q112	[REDACTED]
March 5th	20K179	[REDACTED]
	11X121	[REDACTED]
	02M130	[REDACTED]
March 6th	02M234	[REDACTED]
	11X097	[REDACTED]
	28Q082	[REDACTED]

[REDACTED] is the Assessment Implementation Director at the Queens ISC. He will assist you there.

The Scan Center address is listed below my name. Let me know what time to expect you. Please let us know which of you will be at each of the sites listed above. Thanks

[REDACTED]  
Deputy Executive Director  
Content Assessment and Support and Summative Assessments- Office of  
Accountability Department of Education  
44-36 Vernon Boulevard Room 207  
Long Island City, N.Y. 11101  
tel. [REDACTED]  
fax : [REDACTED]  
email : [REDACTED]

-----Original Message-----

From: [REDACTED]  
Sent: Wednesday, February 27, 2008 7:41 AM  
To: [REDACTED]  
Subject: RE: Audit of DOE's Administration of NYS Standardized Exams

Yes [REDACTED] There will be three teams (1 team for each school).

-----Original Message-----

From: [REDACTED]  
Sent: Tuesday, February 26, 2008 5:30 PM  
To: [REDACTED]  
Subject: Re: Audit of DOE's Administration of NYS Standardized Exams

March 4 I will arrange tomorrow. March 5 and 6 do you want to visit all three schools each day. That is possible to arrange

-----  
Sent from my BlackBerry Wireless Handheld

----- Original Message -----  
From: [REDACTED]  
To: [REDACTED]  
Sent: Tue Feb 26 16:01:10 2008  
Subject: RE: Audit of DOE's Administration of NYS Standardized Exams

Tuesday March 4, 2008

PS 112 (Queens)  
ISC  
SC

Wednesday March 5, 2008

PS 179 (Brooklyn)  
PS 121 (Bronx)  
PS 130 (Manhattan)

Thursday March 6, 2008

PS 234 (Manhattan)  
PS 097 (Bronx)  
PS 082 (Queens)

-----Original Message-----  
From: [REDACTED]  
Sent: Tuesday, February 26, 2008 2:28 PM  
To: [REDACTED]  
Subject: Re: Audit of DOE's Administration of NYS Standardized Exams

Ok. So let me know the schools

-----  
Sent from my BlackBerry Wireless Handheld

----- Original Message -----  
From: [REDACTED]  
To: [REDACTED]  
Sent: Tue Feb 26 14:04:14 2008  
Subject: RE: Audit of DOE's Administration of NYS Standardized Exams

On Tuesday March 4 we would like to go to a school, then accompany the individuals with the answer sheets to the ISC and then proceed to the Scan Centre. The visit to the ISC and the Centre would be for Tuesday only.

----- Original Message -----

From: [REDACTED]  
To: [REDACTED]  
Sent: Tue Feb 26 13:28:16 2008  
Subject: RE: Audit of DOE's Administration of NYS Standardized Exams

Thanks [REDACTED] We will select the schools to visit.

-----Original Message-----

From: [REDACTED]  
Sent: Thursday, February 26, 2008 11:26 AM  
To: [REDACTED]  
Cc: [REDACTED]  
Subject: Re: Audit of DOE's Administration of NYS Standardized Exams

Would you like to select the schools to visit on march 4 5 6 or shall I select? Please let me know

-----  
Sent from my BlackBerry Wireless Handheld

----- Original Message -----

From: [REDACTED]  
To: [REDACTED]  
Cc: [REDACTED]  
Sent: Tue Feb 26 11:10:18 2008  
Subject: Audit of DOE's Administration of NYS Standardized Exams

Good Morning,

This is to inform you that as part of the audit survey stage, we would like to conduct observations of the MATH exam being administered next week (March 4, 5, and 6) in the NYC elementary schools. In addition, we would like to visit the Integrated Service Center in LIC, as well as the Scan Centre on Tuesday, March 4. Please confirm once you have made the necessary arrangements with ISC and Scan Centre officials.

Thank you in advance.

[REDACTED]  
Auditor

\*\*\*\*\*

Sent from the New York City Office of the Comptroller. This email and any files transmitted with it are confidential and intended solely for the use of the individual or entity to whom they are addressed. This footnote also confirms that this email message has been swept for the presence of computer viruses.

\*\*\*\*\*

Appendix H -- Erasure Report

Carla B. Collins, Ph.D.  
1 Larkin Center  
Yonkers, NY 10701  
June 12, 2009

Mr. James Liebman  
New York City Department of Education  
52 Chambers Street  
New York, NY 10007

Dear Mr. Liebman:

I have worked in test administration for over 25 years for public school systems including New York City BOE (1984 - 1995), Chicago Public Schools (1995 - 1997), and Yonkers Public Schools (2008 - 2009) and for a major test publishing vendor, CTB, (1997 - 2008) as its liaison to New York State. Throughout my career, I have reviewed testing materials and documentation and conducted investigations related to potential testing improprieties and have used and reviewed erasure analysis as warranted by the circumstances. As a result of this background and experience, I have substantial familiarity with practices related to the investigation of possible testing improprieties, including erasure analysis. I also have substantial experience and familiarity with testing practices, conditions and behaviors related to the New York State English Language Arts, Mathematics, Science and other New York State tests. Over the years, I have reviewed thousands of test answer sheets with students' bubbled responses.

On June 8, 2009, I was asked by the Office of Assessment at the New York City Department of Education to review two test "bubble" answer sheets for the multiple choice portion of the 2008 New York State fourth grade English Language Arts test. Both answer sheets are discussed on pages 13-14 of a Draft Audit Report on the Administration of New York State Standardized Tests by the New York City Department of Education dated June 2, 2009, number MD08-102A, prepared by The City of New York's Office of the Comptroller. Based on my background as an assessment specialist and my experience with New York State testing practices, behaviors and investigations, including erasure analysis, I was asked to evaluate statements in the Draft Audit Report that the answer sheets of "two students from two different schools" had "an excessive pattern of erasures" or an "excessive number of erasures from incorrect to correct responses," that the erasures on the two answer sheets provided "indicators that may signal testing irregularities," and that the "number of erasures from incorrect to correct responses" -- without any evidence of impropriety within the class or school -- "warrants further investigation."

In order to conduct a thorough evaluation of the student work, it was necessary to review not only the two student answer documents highlighted in the Draft Report but to put those student responses in context by examining 1) the answer documents for the other students in each of the two students' respective classes in the two different schools (an internal validity check), and 2) determine if the student results being questioned were atypical of the student (an external validity check). Based upon this careful review, I conclude that the two students' answer documents do not provide any "indicator that signals testing irregularities" and do not show evidence of cheating or testing irregularities that warrant further investigation, for several reasons.

First, the review of the actual erasures on the two identified answer sheets provides straightforward evidence of a common occurrence in high-stakes testing situations involving separate question and answer sheets. The erasure marks and bubbled answers follow the same answer string (e.g., c-b-a-c) and occur in consecutive questions within one block of the test and therefore strongly indicate that in both instances, the two students skipped questions on the test and accidentally bubbled answers into the wrong row. After almost completing the test and realizing the mistake, the students erased and re-bubbled with the same string of responses (e.g.,



c-b-a-c). The apparent explanation for the erasures is not that there is potential cheating, but that the student moved his or her previously bubbled-in answers from the wrong row to the right row, both right and wrong answers.

Such erasures found on a single student's answer document, in isolation, and without similarities to any other students' erasures in the same class or school are not necessarily considered to be "an excessive pattern of erasures." On a high stakes/high anxiety test, in which hundreds of thousands of students are asked to read questions in a test booklet and bubble in answers on a separate answer sheet, it is a common occurrence for some students throughout different classrooms and schools in New York City and New York State to mistakenly bubble answers onto the wrong row who consequently would need to erase a series of answers/bubbled responses and re-bubble them a row or two up or down from the original. Randomly, or in isolation, these types of erasures by individual students in separate classrooms and schools are acceptable and do not provide evidence of wrongdoing by students, test proctors or others.

Secondly, a potentially suspicious pattern of erasures would be identified when similar erasures or frequency of erasures appeared on multiple answer sheets from the same classroom or school. However, my examination of the answer sheets for the two students' classes showed there were no similar patterns of erasures, or original responses, among any of the students for either case.

Furthermore, for these types of exams, students are encouraged to read over their answers and correct mistakes. Erasures by themselves and changes from incorrect to correct answers do not signal potential testing irregularities unless there is some other evidence suggesting assistance from another student or an adult. Again, examining the other answer sheets from these two students' classes showed no evidence of potential testing improprieties.

Finally, to determine if the test scores in question for the two students were skewed or unexpected, previous achievement levels were questioned. This query revealed that both students did test in a range consistent with previous (2007) and current (2009) test results – with their corrected responses. If, however, their responses, without the erasures/corrections were scored, both would fall far below their typical achievement levels. This comparison of 3 years of achievement data also helps to support the conclusion of no testing irregularity, just the mistake of bubbling in the wrong row and then making the correction.

As a result of this inquiry into the validity of unfair or irregular testing practices, I conclude in my professional judgment, that there was not "an excessive pattern of erasures" or an "excessive number of erasures from incorrect to correct responses," that the erasures on the two answer sheets do not provide "indicators that may signal testing irregularities," and that the "number of erasures from incorrect to correct responses" – without any evidence of impropriety within the class or school – does not "warrant further investigation."

Sincerely,



Carla B. Collins

Attachment - Erasure Report

Student #1																												
Item #	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
Erasure Mark						B	D	C	C	B	[C]	D	A	[D]	C	A	D	C	B	[C]	A	B	B	D				
Final Answer	A	C	D	D	B	<b>C</b>	<b>B</b>	B	D	C	[C]	B	C	[D]	A	D	C	A	D	[C]	B	C	A	B	B	D	A	C

Student #2																												
Item #	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
Erasure Mark								D	C	[C]	B	D	[A]		B	A	D	C	B	[C]	A	B	B	C	A	C		
Final Answer	A	B	C	A	B	C	<b>B</b>	D	[C]	C	B	<b>[A]</b>	D	A	C	B	A	D	[C]	B	C	A	B	B	B	B	A	C

Key:

**Yellow** indicates questions where the student initially misaligned answers bubbled onto the answer sheet by one or two rows, then erased the misaligned answers and re-bubbled the same sequence into the appropriate rows.

**Pink** indicates questions that the student initially skipped and later answered after the misalignment was discovered.

**[Brackets]** indicate a question the student did not need to erase and re-bubble because the misaligned (original) answer and the final answer were the same.

**Blank** indicates no erasure.

**Asterisk\*** indicates the 1 erasure out of the 31 total erasures that doesn't fit the pattern.

**BOLD TYPE** indicates Student 2's incorrect final answers, including three bubbled-in after the student corrected the misalignment.

**Attachment – Erasure Report  
Student Answer Documents**

The following two pages are copies of student answer documents identified in the Comptroller's Draft Report on pages 14-15.

All handwritten markings were added by the audit field staff during the audit fieldwork.


- A circle indicates an erasure mark.
- A check mark in the margin indicates an answer that was changed and resulted in a correct answer.
- An x in the margin indicates an answer that was changed and resulted in an incorrect answer.

Appendix H - Student #1 Answer Document

# NY State Grade 4 English Language Arts Test 2008

**MARKING INSTRUCTIONS**

Make Heavy BLACK marks. Erase cleanly. Make NO stray marks.



**CORRECT MARK:** ●

**INCORRECT MARKS:** ○ ⊗ ⊙ ⊚

*16 Erasures*  
16 = correct

Absent BOOK 1

**BOOK 1**

1.	<input type="radio"/>	<input type="radio"/> B	<input type="radio"/> C	<input type="radio"/> D	15.	<input type="radio"/>	<input type="radio"/> B	<input checked="" type="radio"/> C	<input type="radio"/> D ✓
2.	<input type="radio"/> A	<input type="radio"/> B	<input checked="" type="radio"/> C	<input type="radio"/> D	16.	<input checked="" type="radio"/> A	<input type="radio"/> B	<input type="radio"/> C	<input checked="" type="radio"/> D ✓
3.	<input type="radio"/> A	<input type="radio"/> B	<input type="radio"/> C	<input checked="" type="radio"/> D	17.	<input type="radio"/> A	<input type="radio"/> B	<input checked="" type="radio"/> C	<input checked="" type="radio"/> D ✓
4.	<input type="radio"/> A	<input type="radio"/> B	<input type="radio"/> C	<input checked="" type="radio"/> D	18.	<input checked="" type="radio"/> A	<input type="radio"/> B	<input checked="" type="radio"/> C	<input checked="" type="radio"/> D ✓
5.	<input type="radio"/> A	<input checked="" type="radio"/> B	<input type="radio"/> C	<input type="radio"/> D	19.	<input type="radio"/> A	<input checked="" type="radio"/> B	<input type="radio"/> C	<input checked="" type="radio"/> D ✓
6.	<input type="radio"/> A	<input checked="" type="radio"/> B	<input checked="" type="radio"/> C	<input type="radio"/> D ✓	20.	<input type="radio"/> A	<input type="radio"/> B	<input checked="" type="radio"/> C	<input type="radio"/> D
7.	<input type="radio"/> A	<input checked="" type="radio"/> B	<input type="radio"/> C	<input checked="" type="radio"/> D ✓	21.	<input checked="" type="radio"/> A	<input checked="" type="radio"/> B	<input type="radio"/> C	<input checked="" type="radio"/> D ✓
8.	<input type="radio"/> A	<input checked="" type="radio"/> B	<input checked="" type="radio"/> C	<input checked="" type="radio"/> D ✓	22.	<input type="radio"/> A	<input checked="" type="radio"/> B	<input checked="" type="radio"/> C	<input checked="" type="radio"/> D ✓
9.	<input type="radio"/> A	<input type="radio"/> B	<input checked="" type="radio"/> C	<input checked="" type="radio"/> D ✓	23.	<input checked="" type="radio"/> A	<input checked="" type="radio"/> B	<input type="radio"/> C	<input checked="" type="radio"/> D ✓
10.	<input type="radio"/> A	<input checked="" type="radio"/> B	<input checked="" type="radio"/> C	<input checked="" type="radio"/> D ✓	24.	<input type="radio"/> A	<input checked="" type="radio"/> B	<input type="radio"/> C	<input checked="" type="radio"/> D ✓
11.	<input type="radio"/> A	<input type="radio"/> B	<input checked="" type="radio"/> C	<input type="radio"/> D	25.	<input type="radio"/> A	<input checked="" type="radio"/> B	<input type="radio"/> C	<input type="radio"/> D
12.	<input type="radio"/> A	<input checked="" type="radio"/> B	<input type="radio"/> C	<input checked="" type="radio"/> D ✓	26.	<input type="radio"/> A	<input type="radio"/> B	<input type="radio"/> C	<input checked="" type="radio"/> D ✓
13.	<input checked="" type="radio"/> A	<input type="radio"/> B	<input checked="" type="radio"/> C	<input checked="" type="radio"/> D ✓	27.	<input checked="" type="radio"/> A	<input type="radio"/> B	<input type="radio"/> C	<input type="radio"/> D
14.	<input type="radio"/> A	<input type="radio"/> B	<input type="radio"/> C	<input checked="" type="radio"/> D	28.	<input type="radio"/> A	<input type="radio"/> B	<input checked="" type="radio"/> C	<input type="radio"/> D


*F-30<sup>5</sup>/15*

Appendix H - Student #2 Answer Document

NY State Grade 4 English Language Arts Test 2008

**MARKING INSTRUCTIONS**

Make Heavy BLACK marks. Erase cleanly. Make NO stray marks.



**CORRECT MARK:** ●

**INCORRECT MARKS:** ○ ⊗ ⊖ ⊕

*15 erased*  

---

*12 correct*

○ Absent BOOK 1

**BOOK 1**

1.	●	B	C	D	15.	●	⊗	C	D	✓
2.	A	●	C	D	16.	⊗	B	●	D	X
3.	A	B	●	D	17.	A	●	C	⊗	X
4.	●	B	C	D	18.	●	B	⊗	D	✓
5.	A	●	C	D	19.	A	⊗	C	●	✓
6.	A	B	●	D	20.	A	B	●	D	
7.	A	●	C	D	21.	⊗	●	C	D	✓
8.	A	●	C	⊗	22.	A	⊗	●	D	✓
9.	A	B	⊗	●	23.	●	⊗	C	D	✓
10.	A	B	●	D	24.	A	●	⊗	D	✓
11.	A	⊗	●	D	25.	⊗	●	C	D	✓
12.	A	●	C	⊗	26.	A	●	⊗	D	X
13.	●	B	C	D	27.	●	B	C	D	
14.	A	B	C	●	28.	A	B	●	D	

*Form 81*