

Thomas P. DiNapoli
COMPTROLLER



110 STATE STREET
ALBANY, NEW YORK 12236

STATE OF NEW YORK
OFFICE OF THE STATE COMPTROLLER

August 9, 2011

Dr. John B. King, Jr.
Commissioner
NYS Education Department
State Education Building - Room 111
89 Washington Avenue
Albany, New York 12234

Re: Report 2011-F-6

Dear Commissioner King:

Pursuant to the State Comptroller's authority as set forth in Article V, Section 1 of the State Constitution; and Article II, Section 8 of the State Finance Law, we have followed up on the actions taken by officials of the State Education Department to implement the recommendations contained in our audit report, *Oversight of Scoring Practices on Regents Examinations* (Report 2008-S-151).

Background, Scope and Objective

Regents examinations are statewide tests that are given each year in particular subject areas, such as English, history, mathematics, science, and foreign languages. They provide standardized measures of students' academic performance and schools' effectiveness and adherence to the State's prescribed curricula. The State requires high school students to pass certain Regents exams to earn a high school diploma. Colleges sometimes use students' Regents exam scores in making admission decisions. In addition, exam results are included in the annual report cards the State publishes for each school district, and are taken into account when the academic performance of the districts is evaluated. Regents exams are administered statewide in January, June, and August of each year.

Within the State Education Department (SED), the Office of Assessment Policy, Development and Administration oversees the development and distribution of the Regents exams. Local school officials are responsible for administering the exams and reporting the results to SED. SED issues a Scoring Key and Rating Guide (also known as Rubrics) for each Regents exam. The Rubrics contain the correct answers for the questions with one correct answer (e.g., multiple choice questions) and examples of acceptable answers for the questions with more than one acceptable answer (e.g., fill-in-the-blank questions and essays). The Rubrics also contain guidelines for awarding partial credit where applicable and instructions for converting the "raw" exam score to the final published score.

Regents exams are scored (or “rated”) by teachers at the schools giving the exams. The school principal is responsible for selecting the raters for each exam and monitoring the scoring process to ensure that it is performed in accordance with SED guidelines. Usually, the rater is a teacher who is responsible for teaching the subject covered by the exam. SED has specialists in each examination subject, who may be consulted by the schools during the examination period if it is not clear how a particular question or exam should be scored.

According to SED guidelines, all raters must be thoroughly familiar with the rating instructions for their exams. Training sessions in the scoring process are provided each year by local Boards of Cooperative Educational Services (BOCES) and other designated trainers. Raters may attend these sessions or in-house training sessions at their schools. The raters and other school administrators involved in the scoring process are required to sign a certification stating that they have followed the rules for administering, supervising and scoring the exams. In addition, to ensure scoring accuracy and consistency throughout the State, SED hires expert consultants to analyze scoring variations among schools and individual raters and to periodically perform a statewide review to assess the accuracy of local scoring practices.

Our initial audit report was issued on November 19, 2009. Our objective was to determine whether SED oversight of local school districts provided adequate assurance that the districts accurately scored Regents exams. We found SED reviews of the scoring of selective Regents exams identified significant inaccuracies by local school districts. These inaccuracies tended to inflate the academic performance of students and schools. Although SED detected this problem, it did not adequately ensure that school districts corrected the problem so that future exams would be more accurately scored. We further found that when school districts failed to comply with SED requests to submit scored exams for further review, SED did not follow up to obtain these examinations. In addition, we concluded that SED had limited assurance that exam raters actually attended annual training for scoring exams. The objective of our follow-up was to assess the extent of implementation, as of May 18, 2011, of the twelve recommendations included in our initial report.

Summary Conclusions and Status of Audit Recommendations

Department officials made significant progress in addressing the issues we identified in the initial report. Of the twelve prior audit recommendations, nine have been implemented, two have been partially implemented, and one has been not been implemented.

Follow-up Observations

Recommendation 1

Implement the improvements recommended by the 2003-04 consultant and 2005 Review team reports, or take alternative actions to address the questionable scoring practices they identified.

Status - Implemented

Agency Action - The 2003-04 consultant and 2005 Review team reports recommended improvements in training, scoring calibration, scoring supervision and built-in quality

control during the actual scoring process. In January 2010, SED officials required exam raters to formally certify that they have been trained to score the exam. These certifications are sent to SED. Also in January 2010, SED officials established an acceptable range for exam scoring discrepancies. Officials use this criterion in Department Reviews to evaluate rater scoring consistency through exam re-scoring by independent raters. SED notifies a school when rater discrepancies are outside of the acceptable range. Further, in 2011, SED provided districts with answer keys and training materials on a password protected website and required exam raters to train with the materials before scoring the exams. Also, SED officials advised us that they provided regional information centers (designated BOCES) with power point presentations about proper exam scoring techniques.

Recommendation 2

For each Regents exam that is re-scored in a Department Review, establish an acceptable range for the scoring discrepancies between the Review team and the original raters. Evaluate each school in the sample on the basis of the criteria, and report the evaluations to the schools.

Status - Implemented

Agency Action - In January 2010, SED officials established an acceptable range for scoring discrepancies between the Department review team and the original school rater(s). The discrepancies should be less than two standard deviations from the average score resulting from the Department Review. Using this criterion for the Department Review of the January 2010 Physics Exam, SED officials identified one school that exceeded the acceptable range and subsequently reported the evaluation to that school. Two schools exceeded the acceptable range based on the Department Review of the August 2010 Integrated Algebra Exam, and SED reported its evaluation to one of the schools. At the time of our follow-up review, SED officials were awaiting the results of a formal investigation into the scoring practices of the other school. Officials added they will take actions, as appropriate, pending the results of the investigation.

Recommendation 3

Request schools with significant exam scoring deficiencies to advise SED of any changes made to exam scores as a result of errors identified by the Department Review.

Status - Not Implemented

Agency Action - According to SED officials, the primary purpose of a Department Review is to test scoring consistency and the effectiveness of training provided to achieve such consistency (and not to rescore the exams). Consequently, officials did not implement this recommendation. Officials also noted, however, that schools must report score changes to SED resulting from isolated errors (such as the incorrect addition of points awarded for correct answers) that the schools detect within four months of the test date. Also, if a school administrator believes that widespread scoring errors have occurred, the administrator must formally request and obtain permission from SED to re-score the exams in question.

Recommendation 4

If a school's scoring practices are found to be unacceptable, require the school to develop a corrective action plan and follow up with the school to determine whether the plan is being implemented.

Status - Partially Implemented

Agency Action - SED officials completed investigations of questionable scoring practices at three schools outside of New York City. As a result of the investigations, SED instructed the schools to develop corrective action plans to address certain deficiencies. SED officials believe the schools implemented their corrective action plans. However, officials had no evidence that the schools actually implemented their action plans. SED officials also advised us of ongoing investigations of possible improper scoring practices at four New York City schools. Although the investigations were not complete at the time of our review, SED officials required the schools to implement measures to prevent reoccurrences of the alleged improper practices. SED officials added that corrective action plans will be developed pending the results of the investigations.

Recommendation 5

Obtain and review all examinations that are requested from schools during a Department Review, even if the papers cannot be included in that particular Review.

Status - Partially Implemented

Agency Action - For the Department Review of the January 2010 Physics Exam, SED officials selected 53 schools, and 42 of them submitted their exams in time for inclusion in the Review process. Officials requested and obtained the exams from the remaining 11 schools after the Review was completed. For the Department Review of the August 2010 Integrated Algebra Exam, SED officials selected 71 schools and obtained exams from 67 of them. Officials later learned that two of the four schools did not administer the Algebra Exam in August 2010, and therefore, the two schools had no exam papers to submit for review. The other two schools, however, had administered the Algebra Exam, but did not submit the required papers to SED. Officials added that the schools whose exams were not submitted in time for the 2010 Department Reviews will be included in Reviews scheduled for 2011.

Recommendation 6

Modify the process for selecting schools in Department Reviews to ensure that all schools with a significant presence in the Regents examination program are selected for review within a reasonable period of time.

Status - Implemented

Agency Action - To select schools for the Department Review of January 2010 Physics Exam, SED officials stratified all schools who ordered the exam into six categories based on size, location and economic need. SED officials then determined the percentages of exams from each category and applied those percentages to the overall Review sample. Officials believe this methodology ensures adequate representation of all types of schools (from large urban centers to small rural districts), and over time, most schools will be selected for a Review. Officials used a similar process to select schools for the Department Review of the August 2010 Integrated Algebra Exam.

Recommendation 7

Develop a formal process of assessing a school's risk for irregularities in its scoring of Regents examinations, use this process to assign a level of risk to all the schools in the Regents examination program, and routinely include a certain number of high-risk schools in each Department Review.

Status - Implemented

Agency Action - SED officials focus their attention on exams with scores ranging from 60 to 69 and analyze data to identify schools with irregular score distributions near the minimum passing score of 65. The analysis identifies schools that deviate significantly from the State norms in the percentages of students scoring in the 60-64 and the 65-69 ranges. Using exam data from 2007 through 2010, SED officials assessed schools Statewide and assigned an associated risk level (low, medium or high) to each of them. Officials further indicated that high risk schools are included in all Department Reviews.

Recommendation 8

Perform Department Reviews annually.

Status - Implemented

Agency Action - Since our initial report was issued (November 2009), SED performed Department Reviews of the January 2010 Physics and August 2010 Integrated Algebra exams. Also, the Department plans to conduct Reviews in 2011 and will perform them annually thereafter.

Recommendation 9

Expedite the completion of the January 2008 Department Review.

Status - Implemented

Agency Action - In May 2009, SED completed the Department Review of the January 2008 Regents Examination in English and issued the Review's technical report.

Recommendation 10

Ensure that the examination complaint log is kept up-to-date and accurate, and maintain documentation of all investigations of complaints about examination scoring practices outside New York City.

Status - Implemented

Agency Action - The Office of Assessment Policy Development and Administration maintains an electronic examination complaint log. In April 2011, SED officials added a status screen to the log, allowing staff to check the status of investigations into complaints. We determined that the log was up-to-date (with entries through March 2011) at the time of our follow-up review. We also randomly selected four complaints from the log and verified that Office officials maintained documentation of their investigations into these complaints.

Recommendation 11

Advise school districts to maintain documentation of their raters' attendance at training sessions for Regents exam scoring.

Status - Implemented

Agency Action - Effective January 2010, SED formally required exam raters to certify that they attended exam scoring training sessions. SED further requires school districts to gather the certifications and submit them to SED as evidence of raters' attendance at the sessions.

Recommendation 12

Expand the formal rater exam certification to include an affirmation that the rater attended training for exam scoring.

Status - Implemented

Agency Action - In conjunction with the actions taken to address Recommendation 11, SED expanded the rater certification form to include an affirmation that a rater attended the training for the exam(s) the rater scored.

Major contributors to this report were Karen Bogucki and Donald Collins.

We would appreciate your response to this report within 30 days, indicating any actions planned to address the unresolved issues discussed in this report. We also thank the management and staff of the State Education Department for the courtesies and cooperation extended to our auditors during this review.

Very truly yours,

Brian E. Mason
Audit Director

cc: Ms. Valerie Grey, SED
Mr. James Conway, SED
Mr. Thomas Lukacs, DoB