
Tennessee Department of Education Controls Over State Assessment Scoring

FINAL AUDIT REPORT



ED-OIG/A02I0034
May 2009

Our mission is to promote the efficiency, effectiveness, and integrity of the Department's programs and operations.



U.S. Department of Education
Office of Inspector General
New York, New York

NOTICE

Statements that managerial practices need improvements, as well as other conclusions and recommendations in this report represent the opinions of the Office of Inspector General. Determinations of corrective action to be taken will be made by the appropriate Department of Education officials.

In accordance with Freedom of Information Act (5 U.S.C. § 552), reports issued by the Office of Inspector General are available to members of the press and general public to the extent information contained therein is not subject to exemptions in the Act.



**UNITED STATES DEPARTMENT OF EDUCATION
OFFICE OF INSPECTOR GENERAL**

New York Audit Region
32 Old Slip
Financial Square
New York, New York 10005
Phone: (646) 428-3860 Fax: (646) 428-3868



May 28, 2009

Dr. Timothy K. Webb
Commissioner of Education
Tennessee Department of Education
Andrew Johnson Tower, 6th Floor
710 James Robertson Parkway
Nashville, TN 37243

Dear Dr. Webb:

Enclosed is our final audit report, Control Number ED-OIG/A02I0034, entitled *Tennessee Department of Education Controls Over State Assessment Scoring*. This report incorporates the comments you provided in response to the draft report. If you have any additional comments or information that you believe may have a bearing on the resolution of this audit, you should send them directly to the following Department of Education officials, who will consider them before taking final Departmental action on this audit:

Joseph Conaty
Executive Administrator Delegated the Authority to Perform
the Functions and Duties of the Assistant Secretary for
Office of Elementary and Secondary Education
400 Maryland Ave., SW
Room 3E314
Washington, DC 20202

Carmel Martin
Assistant Secretary
Office of Planning, Evaluation and Policy Development
400 Maryland Ave., SW
Room 7W200
Washington, DC 20202

It is the policy of the U. S. Department of Education to expedite the resolution of audits by initiating timely action on the findings and recommendations contained therein. Therefore, receipt of your comments within 30 days would be appreciated.

In accordance with the Freedom of Information Act (5 U.S.C. § 552), reports issued by the Office of Inspector General are available to members of the press and general public to the extent information contained therein is not subject to exemptions in the Act.

Sincerely,

Daniel Schultz
Regional Inspector General for Audit

Enclosure

List of Selected Acronyms

AYP	Adequate Yearly Progress
BIG	Business Information Group
CTB	CTB/McGraw Hill
EOC	End of Course Tests
ESEA	Elementary and Secondary Education Act, as amended by the No Child Left Behind Act of 2001
FERPA	Family Educational Rights and Privacy Act
FPCO	U.S. Department of Education, Family Policy and Compliance Office
GAO	Government Accountability Office
GIS	Group Information Sheet
LEA	Local Education Agency
MI	Measurement Incorporated
OIG	Office of Inspector General
PEM	Pearson Educational Measurement
PII	Personally Identifiable Information
SAS	SAS Institute Incorporated
SEA	State Educational Authority
SGL	School Group List
TCAP	Tennessee Comprehensive Assessment Program
TDOE	Tennessee Department of Education
TRICOR	Tennessee Rehabilitative Initiative in Correction
TVAAS	Tennessee Value-Added Assessment System

TABLE OF CONTENTS

	<u>Page</u>
EXECUTIVE SUMMARY	1
BACKGROUND	3
AUDIT RESULTS	7
FINDING NO. 1 – Insufficient Monitoring of Contractor Activities for the Tennessee Comprehensive Assessment Program Writing Assessment	7
FINDING NO. 2 – Lack of Written Policies and Procedures and Insufficient Contract Terms.....	12
FINDING NO. 3 – Improper Disclosure of Personally Identifiable Information.....	16
OBJECTIVE, SCOPE, AND METHODOLOGY	20
Enclosure 1: Glossary	22
Enclosure 2: TDOE Comments	23

EXECUTIVE SUMMARY

The objective of our audit was to determine whether controls over scoring of assessments at the Tennessee Department of Education (TDOE) were adequate to provide reasonable assurance that assessment results are reliable. Our review covered assessments administered in school year 2007-2008¹ used for evaluating individual students and making adequate yearly progress (AYP) determinations under Section 1111(b)(3) of the Elementary and Secondary Education Act, as amended by the No Child Left Behind Act of 2001 (ESEA). The importance of valid and reliable assessments is highlighted on the U. S. Department of Education (Department) website for State Fiscal Stabilization Fund requirements. In particular, “. . . the [Department] will award governors approximately \$48.6 billion by formula under the [State Fiscal Stabilization Fund] program in exchange for a commitment to advance essential education reforms to benefit students from early learning through post-secondary education, including . . . high quality, valid and reliable assessments for all students. . . .”

TDOE’s controls over scoring of State assessments were adequate to provide reasonable assurance that assessment results are reliable, except for concerns in three areas. These areas are insufficient monitoring of contractor activities for the Tennessee Comprehensive Assessment Program (TCAP) Writing assessment, lack of written policies and procedures, and insufficient contract terms. Also, TDOE’s disclosure of personally identifiable student information to assessment contractors raises concerns about its compliance with FERPA. Based on our findings, we recommend that the Acting Assistant Secretary for Elementary and Secondary Education require TDOE to:

- 1.1 Review qualifications for a sample of readers to verify that readers are pre-qualified to participate in the live scoring of TCAP Writing assessment based on contract requirements;
- 1.2 Document procedures performed to verify that readers are qualified to participate in scoring;
- 1.3 Strengthen its read-behind process to ensure that readers are scoring TCAP Writing assessment responses in accordance with the approved rubric and anchor sets (see Enclosure 1) by conducting blind reviews on a sufficient sample of scored student responses;
- 1.4 Establish a monitoring plan to ensure that contractors meet contract provisions;
- 1.5 Perform and document monitoring in accordance with the monitoring plan;
- 2.1 Create and distribute written policies and procedures prescribing internal controls for assessment scoring and reporting, including standard procedures for inventorying test material; preparing test materials for scanning; scanning at the central scanning facility; monitoring assessment contractors, required sign-offs, and psychometrical analysis and review; tracking requests for the review of individual student scores; tracking errors; and the documenting of such internal control procedures;
- 2.2 Include provisions in its contracts to facilitate Federal audits and reviews and ensure that contractors adhere to specified industry standards;

¹ Assessments administered during the 2007-2008 school year are those assessments used to calculate adequate yearly progress reported for school year 2007-2008. They include the Spring 2008 TCAP Achievement, the Spring 2008 TCAP Writing, and the Gateway assessments administered in Summer 2007, Fall 2007, and Spring 2008.

- 3.1 Use unique identifiers instead of names, social security numbers, and dates of birth on assessment documents; and
- 3.2 Take appropriate steps to properly protect student personally identifiable information (PII) in the assessment process and ensure that the disclosure of student PII to assessment contractors is in compliance with the Family Educational Rights and Privacy Act (FERPA).

We provided a draft of this report to TDOE for review and comments on March 13, 2009.² In TDOE's comments to the draft report, dated April 14, 2009, TDOE generally concurred with our findings. TDOE did not specifically concur or disagree with our recommendations. However, it provided corrective actions to address all recommendations, except for recommendations 1.3, 3.1 and 3.2, where TDOE provided corrective actions partially addressing recommendations 3.1 and 3.2, and no corrective actions to address recommendation 1.3. We changed the amount identified in the background section of the draft report for Federal funding allocated to the TCAP Gateway Assessment based on additional documentation provided by TDOE. However, the amount of Federal funds allocated to the Gateway contract has no affect on the findings and recommendations in the report. Additionally, in relation to TDOE's comments about its currently employed psychometrician, we clarified the statement in Finding 2 describing recent turnover experienced by TDOE. Otherwise, our findings and recommendations remain unchanged.

Except for personally identifiable information (that is, information protected under the Privacy Act of 1974 (5 U.S.C. § 552a)), the entire narrative of TDOE's comments is included in Enclosure 2. All personally identifiable information mentioned in TDOE's comments was replaced with bracketed text.

² Subsequent to the issuance of the draft, we added further clarification in Finding 3 regarding TDOE's compliance with FERPA.

BACKGROUND

ESEA §1111(b)(3) requires States to implement a set of yearly student academic assessments. The assessments are used as the primary means of determining the yearly performance of the State and each of its local educational agencies (LEAs) and schools in enabling all children to meet the State's challenging student academic achievement standards. States must use the assessments to measure the achievement of students against State academic content and student academic achievement standards in mathematics, reading or language arts, and science. ESEA §1111(b)(3)(C)(iii) states that assessments shall be used for purposes for which such assessments are valid and reliable (see Enclosure 1), and consistent with relevant, nationally recognized professional and technical standards. These requirements significantly increased the urgency for States, LEAs, and schools to produce accurate, reliable, high-quality educational data.

The *Standards for Educational and Psychological Testing*³ differentiates between high and low stakes testing based upon the importance of the results for individuals, organizations, and groups. According to the *Standards for Educational and Psychological Testing*,

At the individual level, when significant educational paths or choices of an individual are directly affected by test performance, such as whether a student is promoted or retained at a grade level, graduated, or admitted or placed into a desired program, the test use is said to have high stakes. . . . Testing programs for institutions can have high stakes when aggregate performance of a sample or of the entire population of test takers is used to infer the quality of service provided, and decisions are made about institutional status, rewards, or sanctions based on test results. . . . The higher the stakes associated with a given test use, the more important it is that test-based inferences are supported with strong evidence of technical quality.

Accordingly, State assessments required by ESEA are considered high-stakes for States, LEAs, and schools for the purposes of calculating and reporting AYP. However, based on the use of the results, these assessments may be considered low-stakes for individual students.

TDOE State Assessments

For the 2007 award year,⁴ TDOE was awarded approximately \$7.6 million in ESEA Title VI funds for State assessments and related activities. TDOE administers three assessments to comply with ESEA §1111(b)(3) – TCAP Achievement, TCAP Writing, and Gateway assessments. In Tennessee, students in grades 3 through 8 take tests in math, science, and reading and language arts, known as the TCAP Achievement. In grades 5, 8, and 11, they also take the TCAP Writing assessment. High school students take Gateway exams in algebra, biology, and English. The results of each of these assessments are used to calculate AYP, thus making each assessment high-stakes for States, LEAs, and schools. High school students must

³ The *Standards for Educational and Psychological Testing* (1999) was developed jointly by the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education.

⁴ The 2007 award year for Federal funds is July 1, 2007, through September 30, 2008.

pass all three Gateway exams to graduate, making this assessment high-stakes for individual students as well.

TDOE entered into a \$41.3 million contract with CTB/McGraw Hill (CTB) for its TCAP Achievement for the period May 1, 2003, through November 15, 2008. Based on the contract, CTB was to develop custom TCAP Achievement testing material; score student responses; provide technology services and accurate on-time student data; provide training to support the project and its staff; and provide expertise in technical fields such as psychometrics. Based on the contract summary sheet about \$11.9 million in Federal funds was allocated to this contract. (See TDOE Assessment Contracts Table on page 4.)

TDOE entered into a \$6.5 million contract with Measurement Incorporated (MI) for scoring of its TCAP Writing assessment for the period January 1, 2004, through June 30, 2008. For the TCAP Writing assessment, student response documents were shipped by the school systems⁵ directly to MI. MI was responsible for securing test materials, hiring readers, training readers based on approved rubrics and anchor sets, scoring student responses, and maintaining an acceptable level of inter-rater reliability (see Enclosure 1) with scoring personnel and the State. Based on the contract summary sheet about \$2 million in Federal funds was allocated to this contract. (See TDOE Assessment Contracts Table on page 4.)

For its Gateway assessments, TDOE entered into a \$19 million contract with Pearson Educational Measurement (PEM) for the period August 11, 2004, through September 30, 2009. TDOE contracted with PEM to conduct the administration, scoring, reporting, analysis, and continued item and form development of the TCAP High School End of Course (EOC) tests. Three of the EOC tests, called Gateway assessment, which include math, science, and language arts, were used for AYP determinations and required for high school students to graduate. Based on the contract summary sheet \$19 million in Federal funds was allocated to this contract. (See TDOE Assessment Contracts Table on page 4.)

TDOE also contracted with other vendors to assist in the test administration and scoring process. TDOE contracted with Tennessee Rehabilitative Initiative in Correction (TRICOR) to prepare used test material for scanning and operate a test distribution center for processing, packaging, distributing, and warehousing tests. TDOE contracted with Business Information Group (BIG) for technical computer consulting and programming services related to student test scanning, test scoring, and report dissemination. BIG provided technical support for the TCAP Achievement, Gateway, EOC, and Competency tests and for software development. Finally, SAS Institute Inc. (SAS) was contracted to provide support for the Tennessee Value-Added Assessment System (TVAAS).⁶ However, the contract was later amended to include AYP calculations.

According to contract summary sheets for each of these six contracts, TDOE's contracts for assessments and other assessment services totaled about \$89.6 million.⁷ About \$43.9 million in Federal funding was allocated to these six contracts in total. (See TDOE Assessment Contracts Table on page 4.)

⁵ TDOE refers to school districts as school systems.

⁶ Through TVAAS, the state of Tennessee has generated a database of longitudinally merged student test data from 1991 to the present. This database enables the examination of the effectiveness of individual districts in Tennessee over a longer historical span than is provided in annual reporting.

⁷ Information about TDOE's contracts reflects amendments executed prior to our January 15, 2009, exit conference.

TDOE Assessment Contracts⁸			
Contractor	Services Provided	Federal Funding⁹	Total Contract Amount
CTB	TCAP Achievement	\$11,917,000	\$41,290,000
PEM	TCAP Gateway	15,744,000	19,044,000
MI	TCAP Writing	2,086,000	6,456,000
SAS	TVAAS Support	2,005,000	8,531,000
BIG	Technical Computer Consulting	8,866,000	8,866,000
TRICOR	Test Distribution Center	0	5,370,000
Totals		\$40,618,000	\$89,557,000

Elements of TDOE’s Scoring Process

TDOE is unique compared to most States in that it scanned its own multiple choice tests, including TCAP Achievement and Gateway assessments. For the TCAP Achievement assessment, TDOE used the services of maximum security prison inmates through TRICOR, to check counts of test material and prepare student response documents for scanning at TDOE’s main scanning facility in Nashville, Tennessee. Its Gateway assessments were scanned by TDOE personnel at regional scanning facilities.¹⁰ Actual scoring was done by the contractors after scanned data were exported to the contractors’ systems. Prior to live scanning, test decks were processed, through a coordinated effort by TDOE and its assessment contractors, to ensure that all scanning, editing, scoring, and reporting functions were working properly.

TDOE had several controls imbedded into its scoring process to ensure accurate and reliable reporting for its Achievement and Gateway assessments. TDOE conducted checks both before and after scoring to ensure accuracy of the answer keys. School system officials were provided a window of opportunity to verify the accuracy of demographic data and to review individual student performance. School systems could also request a State review of an individual student’s score before the data were finalized. When requested, TDOE officials pulled the student’s actual answer sheet and rescored it by hand. In addition, the State’s psychometrician (see Enclosure 1) reviewed and approved record layouts and cut scores (see Enclosure 1) provided by assessment contractors. The psychometrician and SAS conducted an analysis to check the accuracy of equating (see Enclosure 1) and, for its Achievement assessment, reviewed the impact data analysis. Results were compared and discussed with the assessment contractors. For the Gateway assessments the State’s psychometrician conducted an analysis on a sample of raw score data before quick score (see Enclosure 1) reports were issued.¹¹

TDOE relied heavily on its assessment contractor, MI, for scoring the TCAP Writing assessment. After the TCAP Writing assessment was administered to students, MI provided TDOE with a sample of 2,000 student responses for each grade. A portion of the sample was presented to a

⁸ All dollar amounts rounded to the nearest \$1,000.

⁹ According to the contract summary sheets, Federal funding allocated to these contracts included ESEA Title VI and Individuals with Disabilities Education Act funds.

¹⁰ TDOE had nine regional scanning facilities throughout the state, plus additional sites in Hamilton and Memphis.

¹¹ High school students need to perform proficiently on the Gateway assessments to graduate. Quick scores, the actual number of questions answered correctly by a student, were used to measure each student’s performance on the Gateway assessments.

committee of teachers from across the State to establish anchor sets. Anchor sets and scoring rubrics, which were approved by TDOE, were to be used by MI to train readers and to score student responses. The State contract required that readers hired by MI have at least a bachelor's degree and perform sufficiently on qualifying sets to participate in live scoring. After readers were trained, each student response was scored blind (see Enclosure 1) by two different MI readers during live scoring. According to TDOE's contract with MI for the TCAP Writing assessment, MI room directors and readers were required to "maintain a 95% inter-rater reliability rate both with the scoring personnel and with the State." MI tracked inter-rater reliability between its readers and provided TDOE with inter-rater reliability reports and longitudinal score distribution reports (see Enclosure 1) daily. TDOE also conducted read-behinds on a sample of student responses provided by MI each week to calculate inter-rater reliability with the State. Further, MI was required to track reader performance on validity sets circulated during live scoring to verify that readers were properly calibrated throughout the scoring process. TDOE's daily monitoring during operational scoring focused on the distribution of scores awarded by MI readers.

AUDIT RESULTS

TDOE's controls over scoring of State assessments were adequate to provide reasonable assurance that assessment results are reliable, except for concerns in three areas. These areas are insufficient monitoring of contractor activities for the TCAP Writing assessment, lack of written policies and procedures, and insufficient contract terms. Also, TDOE's disclosure of personally identifiable student information to assessment contractors raises concerns about its compliance with FERPA.

FINDING NO. 1 – Insufficient Monitoring of Contractor Activities for the Tennessee Comprehensive Assessment Program Writing Assessment

TDOE's monitoring of MI's activities related to the scoring of the TCAP Writing assessment was not sufficient. TDOE did not adequately monitor MI to ensure that readers were qualified and that reader qualifications were properly documented. Also, TDOE's read-behind process was inadequate to ensure that MI maintained an appropriate level of inter-rater reliability with the State. Without adequate State monitoring, TDOE may not have sufficient assurance that MI complied with the contract and that test scores were reliable.

Lack of Review of Reader Qualifications

Although TDOE specified reader requirements in its contract with MI, it neither monitored all aspects of these requirements nor documented monitoring procedures performed to ensure that MI's readers met the contract qualification requirements. TDOE's contract stipulated that readers must, at a minimum, have a bachelor's degree, submit a writing sample, and perform sufficiently on qualifying sets before they participate in live scoring. However, TDOE did not obtain or review readers' resumes, transcripts, or writing samples to verify that readers used to score the TCAP Writing assessment met the reader requirements stipulated in the contract. Also, TDOE was unable to provide supporting documentation for procedures it performed prior to live scoring to ensure that readers were qualified to score the TCAP Writing assessment based on qualifying set scores.

We initially requested the names and resumes of readers who participated in the scoring of the TCAP Writing assessment from TDOE. TDOE did not monitor the selection of readers and did not have the requested information; therefore, TDOE had to ask MI to provide the data. TDOE stated that to meet our request, MI created a Microsoft Excel file. We found the file to be inadequate for the purposes of our audit because the file did not include reader names or the actual resumes. The file included only information regarding team assignments, reader numbers, employee file numbers, degrees, and qualifying set scores. In response to our second request, MI refused to provide adequate source documentation to TDOE and the Office of Inspector General (OIG). We considered the requested information essential to verify that MI was compliant with contract provisions related to reader qualifications. Accordingly, OIG issued a subpoena requiring that MI provide its written internal policies and procedures for processing, scoring, and reporting; a roster of reader names; and resumes, transcripts, and qualifying set score sheets for all readers who participated in the scoring of the Spring 2008 TCAP Writing assessment.

In response to the subpoena, MI was unable to provide adequate documentation that its readers obtained bachelor degrees for 23 of the 249 readers who participated in training to score the Spring 2008 TCAP Writing assessment.¹² Of the 23 unsupported readers, MI provided expired teacher licenses for 13 readers, indicated that the bachelor degrees for 7 readers were not properly filed and could not be located, provided illegible documents for 2 readers, and could provide only a high school diploma for 1 reader. We were unable to verify that bachelor degrees were required to be a licensed teacher, as the licenses were inactive and dated back to 1968. Therefore, we determined that expired teacher licenses alone were inadequate to support that the 13 readers had obtained bachelor degrees. We noted that the evaluation interview form, which documented MI's evaluation of an individual's ability to score, indicated that two individuals were recommended for scoring only math. In addition, it was specifically recommended that one of these two individuals not score writing. Despite this, these individuals participated in the holistic scoring (see Enclosure 1) of the Spring 2008 TCAP Writing assessment.

Also, the evaluation interview forms for several other readers indicated that close supervision by team leaders was necessary. However, MI could not demonstrate that team leaders provided additional supervision. Although TDOE's contract with MI indicated that team leaders were required to read about five percent of each of their assigned readers' output, MI officials indicated that team leaders were not required to document this procedure. Therefore, MI could not provide documentation that readers were properly supervised by team leaders during live scoring.

In addition, TDOE's contract required MI to retain "books, records, and documents . . . insofar as they related to work performed or money received . . . for a period of three full years from the date of the final payment." However, MI indicated that it could not provide qualifying set score sheets because these documents had been discarded. TDOE indicated that it inspected qualifying set score sheets but did not maintain documentation supporting that this procedure was performed. Thus, TDOE could not support that readers were qualified to score based on qualifying set scores. TDOE did not adequately monitor MI to ensure readers were qualified, reader qualifications were properly documented, and records retained. As a result, TDOE did not have reasonable assurance that the TCAP Writing assessment was scored by qualified readers.

34 C.F.R. §80.40 states "Grantees are responsible for managing the day-to-day operations of grant and subgrant supported activities. Grantees must monitor grant and subgrant supported activities to assure compliance with applicable Federal requirements and that performance goals are being achieved. Grantee monitoring must cover each program, function or activity." Further, 34 C.F.R. §80.42 requires that all financial and programmatic records, supporting documents, statistical records, and other records reasonably considered as pertinent to program regulations be maintained for a period of 3 years.

In addition, The Council of Chief State School Officers' *Quality Control Checklist for Processing, Scoring, and Reporting* indicates that before hand-scoring of open-ended items occurs, States should review readers' qualifications to ensure that the training of scorers is conducted appropriately. Further, standard 3.23 of the *Standards for Educational and*

¹²Of the 249 readers who participated in training, 3 did not qualify according to MI. In total, 246 readers participated in live scoring for the Spring 2008 TCAP Writing.

Psychological Testing indicates that, “the process for selecting, training, and qualifying scorers should be documented by the test developer.”

TDOE indicated that it relied on MI to hire qualified readers and trusted that MI’s hiring practices would comply with contract provisions. TDOE included provisions in its contract with MI, requiring the implementation of specified controls over scoring and scoring personnel. Although TDOE indicated that it monitored other areas of the contract, it did not verify that readers met minimum requirements specified in the contract. TDOE should have monitored assessment contractors to ensure compliance with contract provisions and should have verified that readers were adequately qualified to score. Without adequate State monitoring of reader qualifications, TDOE may not have sufficient assurance that MI complied with the contract and that test scores were reliable.

Read-Behind Process Inadequate

The contract with MI required that at least 2,400 student response documents be selected for State read-behinds. At a minimum MI was required to provide a sample of 200 student response documents per grade during each week of operational scoring, so that TDOE could perform read-behinds to ensure inter-rater reliability with the State. According to daily inter-rater reliability reports and score distribution reports obtained directly from the MI website, scoring began on March 18, 2008, and was completed on April 11, 2008. This represents approximately four weeks of scoring.

MI provided a sample of only 1,080 documents which represent less than one-half of the number required under the contract. In addition, TDOE indicated that read-behinds were conducted on all 1,080 student response documents but provided support for only 990. Thus, MI did not meet the sampling requirements for State read-behinds stipulated in the contract. Further, TDOE did not perform adequate monitoring of MI, because it did not ensure that it received and reviewed an appropriate number of student response documents.

In addition, the number of read-behinds conducted by TDOE may not have been sufficient to properly determine inter-rater reliability with the State. MI was responsible for the reading/scoring of written responses for over 200,000 5th, 8th, and 11th grade students. TDOE’s contract with MI stipulated that readers must maintain a 95 percent inter-rater reliability rate with the scoring personnel and the State. TDOE stated that MI selected a random sample of 1,080 student responses (over all three grades tested) that TDOE officials used to perform read-behinds. However, TDOE was able to support that read-behinds were conducted on only 990 student responses. Furthermore, the number of read-behinds conducted by TDOE was significantly less than the 10 percent of each scorer’s work suggested by the Department’s non-regulatory guidance entitled *Improving Data Quality for Title I Standards, Assessments, and Accountability Reporting: Guidelines for States, LEAs and Schools*.

Finally, the judgment of individuals conducting read-behinds may be tainted by advanced knowledge of initial scores. MI selected a sample of student responses that TDOE used to perform read-behinds. Along with the student response document, the individual conducting the read-behinds received monitor sheets, which recorded the initial scores awarded by MI’s two independent initial readers. Because MI’s scores were provided to TDOE officials and consultants when they conducted read-behinds, the effectiveness of this quality control procedure

was weakened. Advance knowledge of the initial scores may have tainted the judgment of the individual conducting the read-behind activity.

34 C.F.R. §80.40 states “Grantees are responsible for managing the day-to-day operations of grant and subgrant supported activities. Grantees must monitor grant and subgrant supported activities to assure compliance with applicable Federal requirements and that performance goals are being achieved. Grantee monitoring must cover each program, function or activity.”

Standard 5.9 of the *Standards for Educational and Psychological Testing* indicates that, “When test scoring involves human judgment, scoring rubrics should specify criteria for scoring. Adherence to established scoring criteria should be monitored and checked regularly. Monitoring procedures should be documented.” In addition, the Department acknowledged prior recommendations from the Department’s Office of the Inspector General (OIG) and the Government Accountability Office (GAO) by issuing non-regulatory guidance entitled *Improving Data Quality for Title I Standards, Assessments, and Accountability Reporting: Guidelines for States, LEAs and Schools*. This guidance indicates that for constructed response items, States should “. . . ensure that raters’ scores are consistent both across different raters and within the same rater over time.” In the guidance, the Department also points to controls used by the National Assessment of Educational Progress contractors to identify ways States can monitor the location, status, and scoring accuracy of student assessments in real time to prevent reporting errors. One of these controls suggested States, “have supervisors ‘backread’ scored assessments periodically to ensure individual scorers performed consistently over time. Approximately 10 percent of each scorer’s work is backread during the course of scoring.”

In its TCAP Writing assessment contract with MI, TDOE acknowledged that “hand scoring blind responses with accuracy and consistency is a key component in maintaining the reliability and integrity of the TCAP [Writing assessment] programs.” The contract also required MI to provide 200 scored student response documents per grade each week for State read-behinds. TDOE did not take steps to ensure that MI fully complied with this contract provision. Although TDOE indicated that it monitored other areas of the contract, it did not ensure that a sufficient sample of student responses was provided for State read-behinds. States should monitor assessment contractors to ensure compliance with contract provisions and verify that student responses were scored according to approved rubrics and anchor sets. Without adequate State monitoring, TDOE may not have sufficient assurance that MI complied with the contract and that test scores were reliable.

Recommendations

We recommend that the Assistant Secretary for Elementary and Secondary Education require TDOE to —

- 1.1 Review qualifications for a sample of readers to verify that readers are pre-qualified to participate in the live scoring of TCAP Writing assessment based on contract requirements.
- 1.2 Document procedures performed to verify that readers are qualified to participate in scoring.

- 1.3 Strengthen its read-behind process to ensure that readers are scoring TCAP Writing assessment responses in accordance with the approved rubric and anchor sets by conducting blind reviews on a sufficient sample of scored student responses.
- 1.4 Establish a monitoring plan to ensure that contractors comply with contract provisions.
- 1.5 Perform and document monitoring in accordance with the monitoring plan.

TDOE Comments

TDOE concurred with Finding 1. TDOE did not specifically concur or disagree with the recommendations, but stated it implemented actions to address them, with the exception of recommendation 1.3. TDOE recognized the need to document all monitoring of contractor activities for the TCAP Writing assessment. TDOE indicated that it recently entered into a new contract for the TCAP Writing assessment with MI and has begun the process of documentation and more rigorous, periodic monitoring of all contract provisions.

Regarding recommendations 1.1 and 1.2, TDOE indicated that, under its new contract, MI will be required to demonstrate that all readers meet qualification requirements outlined in the contract. TDOE will ensure that all readers of the TCAP Writing assessment are fully qualified and meet all of the requirements set forth in the contract by closely monitoring this process. TDOE plans to monitor documentation that readers have appropriate degrees, have submitted an acceptable writing sample, and have performed sufficiently on qualifying sets before participation in the live scoring process. TDOE stated it would document its monitoring process accordingly.

For recommendation 1.3, TDOE recognized the need to strengthen its read-behind process by ensuring that all read-behinds are completed without advance knowledge of initial scores. However, TDOE disputed the number of read-behinds performed that we reported in the draft report. TDOE indicated that we only counted the number of read-behinds conducted by the Associate Director for the TCAP Writing assessment, and did not count those performed by its two expert readers. TDOE stated that one of its expert readers performed approximately 100 read-behinds for grade 5 while the other expert reader performed approximately 100 read-behinds for grade 11. In total, TDOE indicated that it performed 100 read-behinds per week for each of these two grades.

With regards to recommendations 1.4 and 1.5, TDOE indicated that it had begun the process of establishing a more sufficient monitoring plan with the TCAP Writing assessment contractor. The new contract with MI contains provisions requiring quarterly reports from the vendor. TDOE indicated that it will monitor the quarterly reports to confirm that contract obligations are properly carried out and are progressing on schedule. TDOE also stated that it would document its monitoring and disseminate the results to appropriate parties.

OIG Response

We considered TDOE's comments to Finding 1 and the related recommendations, and commend TDOE for the corrective actions it said it implemented. However, regarding TDOE's comments on recommendation 1.3, we feel that it is necessary to reemphasize the importance of conducting a sufficient sample of read-behinds on TCAP Writing assessment responses. TDOE indicated in its comments that we did not count all the read-behinds conducted; however, while we were

onsite, we requested that TDOE provide documentation to support all read-behinds conducted for the Spring 2008 TCAP Writing assessment. We considered all of the documentation provided and notified TDOE that the documentation included support for 990 read-behinds. Furthermore, the Associate Director for the TCAP Writing assessment stated that no other supporting documentation was available. In addition, TDOE indicated that for its Spring 2008 TCAP Writing assessment, the read-behinds for all three grade levels (5, 8, and 11) were conducted by the Associate Director for the TCAP Writing assessment. In its response to the draft report, TDOE did not provide additional documentation to support that more than 990 read-behinds were performed by TDOE for the Spring 2008 TCAP Writing assessment. Further, even if TDOE actually conducted 100 read-behinds per grade weekly, as it indicated in its response, its TCAP Writing assessment contract indicated that MI was required to provide a minimum sample of 200 student response documents per grade during each week of operational scoring. Therefore, MI did not provide a sufficient sample of scored student responses for read-behind, based on contract provisions. The number of read-behinds conducted by TDOE may not have been sufficient to properly determine inter-rater reliability with the State. TDOE should ensure that it conducts read-behinds on a sufficient sample of scored student response documents. Our finding and recommendations remain unchanged.

FINDING NO.2– Lack of Written Policies and Procedures and Insufficient Contract Terms

TDOE did not have adequate written policies and procedures prescribing standardized internal control procedures for its assessment scoring process. Also, TDOE’s contracts with assessment vendors were missing key provisions including clauses providing for Federal audit and requiring adherence to specified industry standards.

Lack of Written Policies and Procedures

TDOE did not have written policies and procedures documenting its internal control procedures for the inventorying of test material, preparation of test materials for scanning, scanning at the central scanning facility, monitoring of assessment contractors, required sign-offs, psychometrical analysis, documenting and tracking of requests for the review of individual student scores, and tracking of errors.

TDOE provided user manuals for computer programs used to process assessments and administration manuals distributed to school systems, and BIG provided written guidance for regional scanning. However, these documents did not contain a description of the control elements over the areas identified above. For instance, after several discussions with TDOE, we learned that the State psychometrician performed several types of analyses to verify the reliability of assessment results. But none of the documents provided described the purpose of the procedures, how and when they should be performed, with whom the results were communicated, or whether sign-off was required. Also, there were no written policy and procedure manuals available at the TRICOR warehouse or the cutting facility at the Tennessee Prison for Women.

In addition, TDOE lacked a process to adequately track requests for reviews of student assessment scores. In Tennessee, school systems may request that an individual student’s score be reviewed prior to the finalization of scoring data. We asked that TDOE provide us with these requests, as well as the results of the reviews for both the 2006-2007, and the 2007-2008 school

years to identify possible scoring errors. Because TDOE did not have written policies and procedures, TDOE employees did not have a structured process to document and track school system requests for the review of individual student scores. TDOE officials provided us with either the emails sent by school system personnel or lists they created specifically to meet our request for the information. According to the information provided, TDOE received a total of 52 requests for a review of individual student scores from school systems for the 2007-2008 school year. TDOE was unable to provide complete information for the Spring 2007 TCAP Achievement assessment as this information could not be located. TDOE did not have a standard process for tracking requests for the review of individual student scores to ensure that all requests were actually addressed and handled consistently. It is important that TDOE properly track and document these requests because school system personnel worked more directly with Tennessee students and were in a better position to identify errors in individual student scores.

The central scanning facility also lacked written policies and procedures. Although the Executive Director for Assessment Evaluation & Research stated that a service log was maintained and stored inside each scanning machine, TDOE was unable to provide the log. TDOE stipulated in its contract with Scantron that “the contractor shall perform preventative maintenance inspections . . . per manufacturer’s recommendation . . .”¹³ Service logs are good tools for monitoring machine maintenance to ensure that the equipment was serviced and properly functioning. TDOE relied on Scantron, the service contractor, to perform proper maintenance. When we asked the Scantron representative to provide the logs for our review, he stated that logs were not maintained. As a result, TDOE could not ensure that its scanning machines were properly serviced and calibrated before and during live scanning of assessments.

Also, TDOE was unable to provide the written policies and procedures followed by its contractors in processing its State assessments. When TDOE’s assessment contractors refused to provide requested documents, we issued subpoenas to CTB, PEM and MI requiring that written policies and procedures for processing, scoring, and reporting the TCAP Achievement, Gateway, and TCAP Writing assessments be provided. Although CTB and PEM provided written policies and procedures, MI indicated that it did not maintain written policies and procedures. MI stated that procedures for processing, scoring, and reporting assessments were developed through communication and experience over time.

Improving Data Quality for Title I Standards, Assessments, and Accountability Reporting: Guidelines for States, LEAs and Schools, Section 4.2 identifies transparency as a general principle under data review and validation. It states that “information on all ongoing data quality monitoring procedures should be collected as a staff resource and archived as a continuing reference. Handbooks on management controls should include descriptions of valid data elements from the data dictionary and processes for correcting errors.”

Standard 5.8 of the *Standards for Educational and Psychological Testing* states “test scoring services should document the procedures that were followed to assure accuracy of scoring. The

¹³ The *Operator’s Guide for the 5000i Scanner* provides a detailed maintenance schedule required to keep the scanner in good working order. TDOE’s central scanning facility used NCS 5000i scanners. However, TDOE indicated that the scanners were recently upgraded to increase their productivity from 7,000 pages per hour to 15,000 pages per hour.

frequency of scoring errors should be monitored and reported to users of the service on reasonable request. Any systematic source of scoring errors should be corrected.”

TDOE administered assessments several years before the implementation of the No Child Left Behind Act of 2001. It relied heavily on its experienced staff members to guide those with less experience on proper procedures in the administration and processing of assessments. Many elements of the assessment process were included in documents such as requests for proposals, contracts with assessment vendors, and assessment administration manuals. However, these documents did not incorporate several significant elements of TDOE’s internal controls over assessment scoring. With no written policies and procedures, new and veteran employees may not have the tools necessary to perform internal control procedures appropriately, consistently, and timely. This situation is heightened by employee turnover. According to a TDOE official, two of three psychometricians left TDOE in recent years. Also, this was the first administration of TDOE assessments experienced by the Assistant Director for Elementary Assessments in this position. Finally, during the course of our audit, the Assistant Director for Secondary Assessments left the position.

Inadequate Contract Terms

TDOE’s contracts were not adequate to facilitate proper Federal audit and review. TDOE had contracts with CTB, PEM, and MI for its TCAP Achievement, Gateway, and TCAP Writing assessments, respectively. It also contracted with BIG for technical computer consulting and programming services and SAS for TVAAS support and adequate yearly progress calculations. None of the contracts included an adequate audit clause to provide for audits by Federal officials despite the fact that TDOE contracted with these vendors to assist in its effort to meet Federal requirements under ESEA §1111. In addition, each contract was paid, at least in part, with Federal funds.

During our audit, we had difficulty in obtaining adequate source documentation from TDOE’s assessment contractors. Although the contracts for assessment services with CTB, PEM, and MI contained clauses that authorized State audits, TDOE was unable to obtain this documentation from its contractors.

Also, neither TDOE’s requests for proposals nor its contracts with MI for its TCAP Writing assessment and CTB for its TCAP Achievement assessment included a clause requiring adherence to specified industry standards.

34 C.F.R. §76.910 states “A grantee or subgrantee shall cooperate with the Secretary and the Comptroller General of the United States or any of their authorized representatives in the conduct of audits authorized by Federal law. This cooperation includes access without unreasonable restrictions to records and personnel of the grantee or subgrantee for the purpose of obtaining relevant information.”

ESEA §1111(b)(3)(C)(iii) states that assessments shall “be used for purposes for which such assessments are valid and reliable, and be consistent with relevant, nationally recognized professional and technical standards.”

Tennessee's State procurement policies and procedures did not require the inclusion of a contract provision for the use of specified industry standards in their assessments. Also, the TDOE officials at the time the contracts were created were unaware that such a provision was necessary.

Because TDOE did not have an adequate Federal audit clause in its contracts with assessment vendors, and three contractors refused to provide requested information to both TDOE and OIG, we had to issue subpoenas to TDOE's three primary assessment contractors to obtain the requested documentation. This delayed the Federal audit. It is also an impediment for TDOE in any planned monitoring activities because the contractors refused to provide TDOE with critical information. If Federal auditors are unable to gain access to records and personnel to obtain relevant information while conducting audits, Federal funds and programs may be at risk.

Also, if assessment contractors are not required to adhere to specified industry standards, the reliability of assessment results could be jeopardized. Further, if a contractor actually did fail to adhere to industry standards, that contractor may avoid penalties if the requirement to adhere to industry standards was not specified as a provision in the contract.

Recommendations

We recommend that the Assistant Secretary for Elementary and Secondary Education require TDOE to—

- 2.1 Create and distribute written policies and procedures prescribing internal controls for assessment scoring and reporting, including standard procedures for inventorying test material; preparing test materials for scanning; scanning at the central scanning facility; monitoring assessment contractors, required sign-offs, psychometrical analysis and review; tracking requests for the review of individual student scores; tracking errors; and the documenting of such internal control procedures.
- 2.2 Include provisions in its contracts to facilitate Federal audits and reviews and ensure that contractors adhere to specified industry standards.

TDOE Comments

TDOE concurred with Finding 2. TDOE did not specifically concur or disagree with the recommendations, but stated it implemented actions to address them. TDOE indicated that it currently has written policies and procedures specific to each assessment program, but acknowledged the need to strengthen the overall procedure for handling updates and/or changes to its current written policies and procedures on a more regular basis.

For recommendation 2.1, TDOE indicated that it has begun an extensive reorganization of its assessment archives with an in-depth inventory of all archived assessment documents and their locations. TDOE indicated that it will begin the process of regularly updating all written policies and procedures to include the most current information and will create a more comprehensive list of yearly assessment activities specific to each assessment program.

With regards to recommendation 2.2, TDOE indicated that it has begun the process of including provisions in its contracts to facilitate Federal audits and reviews.

In addition, TDOE indicated in its additional comments and final thoughts that information included in Finding 2 related to the number of psychometricians was inaccurate. TDOE stated that its currently employed psychometrician had been with the TDOE Office of Assessment for the past 17 years.

OIG Response

We considered TDOE's comments to Finding 2 and the related recommendations, and commend TDOE for the corrective actions it said it already implemented. With regards to recommendation 2.1, TDOE should create and continue to update its written policies and procedures to ensure the inclusion of current information related to standard internal control procedures for test material; preparing test materials for scanning; scanning at the central scanning facility; monitoring assessment contractors, required sign-offs, psychometrical analysis, and review; tracking requests for the review of individual student scores; tracking errors; and documenting implementation of these internal control procedures. We feel that it is necessary to reemphasize this point because, as reported in Finding 2, TDOE did not include these elements in its manuals provided to us during our audit.

Regarding TDOE's additional comments about its psychometrician, the information we reported in our draft report was obtained from interviews with a TDOE official held during our audit. During these interviews, the TDOE official indicated that TDOE previously employed three psychometricians to support assessment activities, but it lost two of these employees. We clarified this point in Finding 2. However, this information does not otherwise change Finding 2 or its related recommendations.

FINDING NO.3–Improper Disclosure of Personally Identifiable Information

TDOE's disclosure of student PII to its assessment contractors raises concerns about its compliance with FERPA. TDOE disclosed computer files and student response documents to its assessment contractors, including CTB, PEM, MI, BIG, TRICOR, and SAS to process State assessments required under ESEA. The computer files and student response documents included student names, dates of birth, social security numbers, and test performance data. However, there was no indication that TDOE obtained prior written consent from parents or met the requirements of an exception as required by FERPA.

In addition, students' PII on TCAP Achievement student response documents were disclosed to maximum security prison inmates through their participation in a program run by TRICOR. The disclosure of this highly sensitive information to maximum security prison inmates is a significant security risk that could lead to the misuse of student PII.

TDOE used personally identifiable student information, including names, social security numbers, and dates of birth on student answer sheets. School systems participating in the TCAP Achievement assessment sent used test materials to the warehouse managed by TRICOR. At the TRICOR warehouse, the actual student answer documents were counted and inventoried by inmates from the Tennessee Prison for Women. Counts performed at the warehouse were compared to school system counts recorded on the Group Information Sheet (GIS) as well as the School Group List (SGL). The student answer documents were then shipped to the cutting facility located at the Tennessee Prison for Women. Again, inmates physically counted the

student answer sheets and verified counts to the GIS and SGL. Inmates also cut the spines of the student answer documents to prepare the documents for scanning.

The use of student social security numbers on assessment documents presents opportunities for misuse. Further, the disclosure of this information to prison inmates significantly increases the risk for identity theft. We consulted with the U.S. Department of Education, Family Policy and Compliance Office (FPCO) about our concerns regarding TDOE's contract with TRICOR, and FPCO staff agreed that this situation presented a FERPA issue.

Per 34 C.F.R. §99.30, the parent or eligible student shall provide a signed and dated written consent before an educational agency or institution discloses PII from the student's education records, except as provided in section 99.31.

34 C.F.R. §99.31 allows schools to disclose those records, without consent, to the following parties or under the following conditions:

- School officials with legitimate educational interest;
- Other schools to which a student intends to transfer;
- Specified officials for audit or evaluation purposes;
- Appropriate parties in connection with financial aid to a student;
- Organizations conducting certain studies for or on behalf of the school;
- Accrediting organizations;
- To comply with a judicial order or lawfully issued subpoena;
- Appropriate officials in cases of health and safety emergencies; and
- State and local authorities, within a juvenile justice system, pursuant to specific State law.

TDOE used student social security numbers to track student performance. TDOE is taking steps to change this and use unique student identifiers instead. In the meantime, the use of student PII on assessment documents that are disclosed to maximum security prison inmates significantly increases the risk for identity theft. The use of TRICOR is mandated by Tennessee State law. It was created to provide occupational and life skills training for Tennessee's incarcerated population through job training, program opportunities, and transitional services designed to assist offenders with a successful reintegration into society.

Also, TDOE indicated that its disclosure of student PII in processing assessments met the exceptions at 34 C.F.R. §99.31 (a)(3), (a)(5)(i), and (a)(6)(i). However, we determined, with the assistance of FPCO, that §99.31 (a)(5)(i) and (a)(6)(i) did not appear to apply and that TDOE may not have met the requirements of §99.31 (a)(3).

Section 99.31 (a)(3) allows educational agencies to disclose education records to authorized representatives of State and local educational authorities, in connection with an audit or evaluation of Federal or State supported education programs or for the enforcement of or compliance with Federal legal requirements related to those programs. A memorandum issued by the Deputy Secretary of Education on January 30, 2003, explained that an "authorized representative" under this exception must be under the direct control of the State educational authority (SEA). To extend this exception to a contractor, the SEA and the contractor must enter into an agreement that establishes the SEA's direct control over the contractor's access to and use of PII from the SEA's education records. Per FPCO, direct control means giving direction to the

contractor and explaining how records are to be handled to ensure FERPA compliance. We recognize that final regulations amending FERPA were recently published. These final regulations now include language that specifically allows an SEA to disclose student PII to contractors under the SEA's direct control.¹⁴

TDOE's agreements with its assessment contractors may not adequately establish TDOE's direct control over the contractors' access to and use of PII from education records. TDOE indicated that within the last year, it began requesting its contractors to sign forms that specifically address FERPA. However, TDOE indicated that its contracts with CTB, PEM, MI, BIG, TRICOR, and SAS originated before TDOE began using these forms. Although TDOE has required assessment vendors currently under contract with TDOE to sign the form, there were no forms on file for the assessment contracts covering our audit period. Therefore, there is no indication that TDOE maintained direct control over any of its assessment contractors. Given this, TDOE's disclosure of PII to assessment contractors raises concerns about its compliance with FERPA. In addition, the disclosure of Tennessee student social security numbers to maximum security prison inmates by TDOE presents a security risk for identity theft.

Recommendations

We recommend that the Assistant Secretary for Elementary and Secondary Education require TDOE to —

- 3.1 Use unique identifiers instead of names, social security numbers, and dates of birth on assessment documents.
- 3.2 Take appropriate steps to properly protect student PII in the assessment process and ensure that the disclosure of student PII to assessment contractors is in compliance with the FERPA.

TDOE Comments

TDOE concurred with Finding 3. TDOE did not specifically concur or disagree with the recommendations, but stated it implemented actions to address them. TDOE stated that it used social security numbers in the past to comply with Tennessee law. However, effective July 1, 2008, this law was changed. TDOE stated it will no longer use student social security numbers as of the 2009-2010 school year.

For recommendation 3.1, TDOE stated that it had already begun the process of moving away from the use of student social security numbers. TDOE indicated that it will transition during the current testing year (2008-2009) by linking social security numbers to unique student identification numbers.

For recommendation 3.2, TDOE indicated that all assessment contracts include a FERPA clause. Additionally, all TDOE assessment vendors must sign an *Authorization to Acknowledgement of Compliance* form agreeing to fully comply with FERPA by maintaining the confidentiality of all student information.

¹⁴ 73 Federal Register 74806 and 74816 (December 9, 2008).

OIG Response

We considered TDOE's comments to Finding 3 and the related recommendations, and commend TDOE for the corrective actions it said it already implemented. However, TDOE's response did not fully address recommendation 3.1. Even without the presence of social security numbers, the use of names and dates of birth on assessment documentation presents opportunities for misuse. This risk is especially high in this situation because TDOE discloses student PII recorded on assessment documents to maximum security prison inmates. TDOE needs to protect all student PII, including student names and dates of birth from improper disclosure that puts that information at risk.

Although TDOE stated that its assessment contracts include FERPA clauses and that assessment vendors are required to sign an *Authorization to Acknowledgment of Compliance* form, TDOE should work with OESE and FPCO to confirm that the use of these methods is sufficient to ensure that the disclosure of student PII to assessment contractors is in compliance with FERPA.

OBJECTIVE, SCOPE, AND METHODOLOGY

The objective of our audit was to determine whether controls over scoring of assessments at the TDOE were adequate to provide reasonable assurance that assessment results are reliable. Our review covered assessments administered in school year 2007-2008 used for evaluating individual students and making AYP determinations under Section 1111(b)(3) of ESEA.

To accomplish our objective for our audit of TDOE, we—

- Compiled and reviewed media reports, prior OIG audit reports and memorandum, prior audit reports issued by the Government Accountability Office and the Tennessee State Comptroller, and the monitoring reports and peer review determination letter issued by the Office of Elementary and Secondary Education (OESE).
- Interviewed officials from OESE and the Counsel of Chief State School Officers, as well as officials from independent research organizations including the National Board on Educational Testing and Public Policy (NBETPP) at Boston College, Education Sector, the National Center for Fair and Open Testing, and Eduventures.
- Identified Title VI and Title I, Part A funding and contacted the Office of Elementary and Secondary Education for additional information on TDOE.
- Reviewed the database of errors compiled by the NBETPP to identify potential risk areas in TDOE's system for scoring State assessments.
- Reviewed laws, regulations, and guidance applicable to State assessment scoring and student privacy rights.
- Obtained an understanding of the assessments administered by TDOE, as well as the control system in place to ensure reliable student scores through interviews of relevant personnel and direct observation of scanning at TDOE's central scanning facility in Nashville and two regional scanning facilities. We also observed operations conducted by inmates of the Tennessee Prison for Women through TRICOR.
- Reviewed requests for proposals, vendor proposals, and contracts for the TCAP Achievement, Gateway, and TCAP Writing assessments. We reviewed documentation supporting TDOE's procurement process in relation to these three contracts.
- Reviewed the contracts with vendors for equipment maintenance, technical computer consulting, TVAAS support, and operation of the test distribution center.
- Reviewed the read-behind process conducted by TDOE, analyzed the MI's inter-rater reliability with the State, and determined that the correct score was carried forward as the final score.
- Reviewed inter-rater reliability reports for MI's initial readers to determine whether MI maintained 95 percent inter-rater reliability daily and cumulatively with scoring personnel.
- Reviewed validity reports and score distribution reports provided by MI to TDOE. We also reviewed the validity set circulation log to ensure that all readers completed validity sets as required by the contract.
- Reviewed a list of requests from school systems for the review of individual student scores. We also asked TDOE officials about current and past scoring errors and related corrective actions.

- Reviewed documentation provided to support monitoring of contractor services.
- Issued subpoenas to CTB and PEM for written policies and procedures, and to MI for reader qualification documentation and written policies and procedures.
- Reviewed written policies and procedures for processing, scoring, and reporting provided by CTB and PEM.
- Reviewed qualification documentation provided by MI for the 249 readers who participated in training to score the Spring 2008 TCAP Writing assessment. From the documentation provided, we verified the degrees obtained from non-foreign educational institutions for a random sample of 21 readers out of a population of 208.¹⁵ We conducted alternative procedures to verify the 15 bachelor degrees obtained from foreign educational institutions.

With the assistance of a Computer Assisted Assessment Technologies team member, we selected a random sample of 50 Gateway student response documents from a population of 194,825, and 50 TCAP Achievement student response documents from a population of 433,108. We then compared the hardcopy student response documents with the electronic data file provided by the contractors to ensure accurate scanning and that the scanned data maintained its integrity during export to the contractor's system. We also recalculated raw scores to ensure that they were calculated accurately using the correct answer key. We did not identify any errors in our sample and determined, with a 90 percent confidence level, that the error rate was less than five percent.

We began our work onsite at TDOE in Nashville, Tennessee, on April 28, 2008. We held an exit conference with TDOE officials on January 15, 2009.

We conducted this performance audit in accordance with generally accepted government auditing standards appropriate to the scope of the review described above. Those standards require that we plan and perform the audit 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.

¹⁵ The population of 208 readers did not include 23 readers for whom MI could not provide adequate support or 15 readers for whom foreign degrees were provided.

Enclosure 1: Glossary

Anchor Sets According to TDOE’s contract with MI, “sample essays which clearly exhibit the criteria for each score point.”

Blind Scoring According to TDOE’s contract with MI, “a method of scoring that allows each student response to be scored independently, without the reader having knowledge of any previous scores.”

Cut Score According to the *Standards for Educational and Psychological Testing*, a specified point on a score scale, such that scores at or above that point are interpreted differently from scores below that point.

Equating According to the *Standards for Educational and Psychological Testing*, places two or more essentially parallel tests on a common scale.

Holistic Scoring According to the *Standards for Educational and Psychological Testing*, a method of obtaining a score on a test, or a test item, based on a judgment of overall performance using specified criteria.

Inter-rater Reliability According to the *Standards for Educational and Psychological Testing*, the consistency with which two or more judges rate the work or performance of test takers.

Inter-rater Reliability Report Reports provided by MI showing inter-rater reliability levels for initial readers both daily and cumulatively.

Longitudinal Score Distribution Report Report provided daily by MI, showing the percentage of completion for each grade level, percentage of students awarded each available score value for the current administration and the prior four test administrations.

Psychometrician According to *Margins of Error: The Education Testing Industry in the No Child Left Behind Era*, a report published by Education Sector, experts trained in measurement theory and statistics who analyze how test items perform in field trials and conduct other statistical analyses in test-making.

Quick Score The actual number of questions answered correctly by a student; used by TDOE to measure each student’s performance on the Gateway assessments.

Reliability According to the *Standards for Educational and Psychological Testing*, the degree to which test scores for a group of test takers are consistent over repeated applications of a measurement procedure and hence are inferred to be dependable, and repeatable.

Rubrics According to TDOE’s contract with MI, “scales (currently six point) used to determine the range of excellence in the papers written in response to a given prompt.”

Validity According to the *Standards for Educational and Psychological Testing*, the degree to which accumulated evidence and theory support specific interpretations of test scores entailed by proposed uses of a test.

Enclosure 2: TDOE Comments



PHIL BREDESEN
GOVERNOR

STATE OF TENNESSEE
DEPARTMENT OF EDUCATION
6th FLOOR, ANDREW JOHNSON TOWER
710 JAMES ROBERTSON PARKWAY
NASHVILLE, TN 37243-0375

TIMOTHY K. WEBB, Ed.D.
COMMISSIONER

April 14, 2009

Mr. Daniel Schultz
Regional Inspector General for Audit
U.S. Department of Education
Office of Inspector General
32 Old Slip
26th Floor, Financial Square
New York, New York 10005

RE: Response to Draft Audit Report Control Number ED-OIG/A02I0034

Dear Mr. Schultz:

Thank you for an opportunity to respond to the draft audit report on *Tennessee Department of Education Controls Over State Assessment Scoring*. Below is our response to the findings and recommendations:

OIG FINDING NO. 1 – Insufficient Monitoring of Contractor Activities for the Tennessee Comprehensive Assessment Program Writing Assessment

TDOE Comments to OIG Finding No. 1:

TDOE concurs with this finding. TDOE recognizes the need to document all monitoring of contractor activities for the Tennessee Comprehensive Assessment Program Writing Assessment. TDOE has recently entered into a new Writing Assessment contract with Measurement, Inc. and has begun the process of documentation and more rigorous, periodic monitoring of all contract provisions.

OIG Recommendations

- 1.1 Review qualifications for a sample of readers to verify that readers are pre-qualified to participate in the live scoring of TCAP Writing assessment based on contract requirements.
- 1.2 Document procedures performed to verify that readers are qualified to participate in scoring.
- 1.3 Strengthen its read-behind process to ensure that readers are scoring TCAP Writing assessment responses in accordance with the approved rubric and anchor sets by conducting blind reviews on a sufficient sample of scored student responses.
- 1.4 Establish a monitoring plan to ensure that contractors comply with contract provisions.
- 1.5 Perform and document monitoring in accordance with the monitoring plan.

Mr. Daniel Schultz
Page 2

TDOE Comments to OIG Recommendations:

With regard to OIG Recommendation 1.1, TDOE did not require Measurement, Inc., the assessment vendor, to provide evidence that they had acquired adequate documentation for readers during the live scoring of the 2008 Writing Assessment. Under the new contract, Measurement Inc. will be required to demonstrate that all readers meet qualifications outlined in the contract requirements. TDOE will ensure that all readers of the TCAP Writing Assessment are fully qualified and meet all of the requirements set forth in the terms of the contract by closely monitoring this process with Measurement, Inc. Monitoring will include documentation that readers have appropriate degrees, have submitted an acceptable writing sample, and have performed sufficiently on qualifying sets of student writing papers before participation in the live scoring process.

With regard to OIG Recommendation 1.2, TDOE will verify the qualifications of all readers of the TCAP Writing Assessment employed by Measurement, Inc. as mentioned in the previous recommendation and will document this process accordingly.

With regard to OIG Recommendation 1.3, it should be noted that the auditors counted only the number of read-behind papers conducted by [EMPLOYEE A], the Associate Director for the TCAP Writing Assessment. This number did not include one hundred papers at grades 5 and 11 that [READER A] and [READER B], expert readers in North Carolina, were reading weekly. TDOE read approximately one hundred papers per grade level weekly. [READER A] read approximately one hundred grade 5 papers weekly. [READER B] read approximately one hundred grade 11 papers weekly. With that said, TDOE recognizes the need to strengthen the read-behind process by ensuring that all read-behinds are completed without advance knowledge of initial scores.

With regard to OIG Recommendation 1.4, TDOE has begun the process of establishing a more sufficient monitoring plan with the TCAP Writing Assessment vendor, Measurement, Inc., to ensure compliance with contract provisions. The new TCAP Writing Assessment contract contains provisions for a required quarterly report from Measurement, Inc.

With regard to OIG Recommendation 1.5, the quarterly reports, required in the TCAP Writing Assessment contract with Measurement, Inc., will be routinely monitored to confirm that all contract obligations are carried out as designed and are progressing on schedule. Results of the monitoring plan will be documented and disseminated to all appropriate parties.

OIG FINDING NO. 2 – Lack of Written Policies and Procedures and Insufficient Contract Terms

TDOE Comments to OIG Finding No. 2:

TDOE concurs with this finding. TDOE currently has written policies and procedures specifically relating to each assessment program; however, TDOE acknowledges the need to strengthen our overall procedure for handling updates and/or changes to these documents on a more regular (yearly) basis.

OIG Recommendations

2.1 Create and distribute written policies and procedures prescribing internal controls for assessment scoring and reporting, including standard procedures for inventorying test material;

Mr. Daniel Schultz
Page 3

preparing test materials for scanning; scanning at the central scanning facility; monitoring assessment contractors, required sign-offs, psychometrical analysis and review; tracking requests for the review of individual student scores; tracking errors; and the documenting of such internal control procedures.

2.2 Include provisions in its contracts to facilitate Federal audits and reviews and ensure that contractors adhere to specified industry standards.

TDOE Comments to OIG Recommendations:

With regard to OIG Recommendation 2.1, it should be noted that TDOE currently has in place procedures and implementation manuals specific to each assessment program that include assessment information; test development; pre-assessment, during-assessment, and post-assessment procedures; test processing; professional development; special projects; services to LEAs; special populations; websites; technologies; resources; and timelines of critical activities. TDOE also currently has processing manuals which include the specific procedures for scanning and editing. Based on this recommendation, TDOE has begun an extensive reorganization of its assessment archives with an in-depth inventory of all archived assessment documents and their locations. TDOE will begin the process of regularly updating all written policies and procedures to include the most current information and will create a more comprehensive list of yearly assessment activities specific to each assessment program. Additionally, the new Achievement contract includes a new method for post processing inventory (bar codes).

With regard to OIG Recommendation 2.2, TDOE has begun the process of including provisions in its contracts to facilitate Federal audits and reviews. The End of Course Assessment Request for Proposals was amended to include these provisions. Additionally, should TDOE get the opportunity to make an amendment to the existing contracts for the Achievement and Writing Assessments, it will immediately do so. It has been noted that any and all future TDOE assessment vendor contracts should and will include these provisions.

OIG FINDING NO. 3 – Improper Disclosure of Personally Identifiable Information

TDOE Comments to OIG Finding No. 3:

TDOE concurs with this finding. TDOE has in the past used student social security numbers in keeping with Tennessee law at the time. TCA 49-6-5102 *Social Security number to be provided to admissions officer* was deleted and replaced with TCA 49-6-5101 *Assignment of personal identification numbers* effective July 1, 2008. TCA 49-6-5101 states “To facilitate the identification and tracking of students from school to school and LEA to LEA, and to facilitate and make more efficient the keeping of records, all students shall be assigned a personal identification number upon entry into a Tennessee public school.” [Acts 1990, ch. 708, § 1; 2007, ch. 306, § 1.] TDOE is currently transitioning from the use of student social security numbers and will no longer require student social security numbers as of the 2009-2010 school year.

OIG Recommendations

3.1 Use unique identifiers instead of names, social security numbers, and dates of birth on assessment documents.

Mr. Daniel Schultz
Page 4

3.2 Take appropriate steps to properly protect student PII in the assessment process and ensure that the disclosure of student PII to assessment contractors is in compliance with the FERPA.

TDOE Comments to OIG Recommendations:

With regard to OIG Recommendation 3.1, TDOE has begun this process by moving away from the use of student social security numbers. TDOE will transition during the current testing year (2008-2009) by linking social security numbers to unique student identification numbers. TDOE will no longer use student social security numbers by the 2009-2010 testing year.

With regard to OIG Recommendation 3.2, all TDOE assessment contracts include a Family Educational Rights and Privacy Act (FERPA) clause. Additionally, all TDOE assessment vendors must sign an *Authorization to Acknowledgement of Compliance* form agreeing to fully comply with FERPA by maintaining the confidentiality of all student information.

Additional TDOE Comments and Final Thoughts:

TDOE wishes to note two changes needed in the draft audit report for TDOE Controls over State Assessment Scoring. In the TDOE Assessment Contracts Table included in the background information, federal funding indicated for the TCAP Gateway is incorrectly listed as \$19,044,000. The correct amount allocated from federal funds for the TCAP Gateway Assessment should be listed as \$15,744,000. Additionally, the supporting information included with Finding No. 2 states that "in recent years, TDOE lost several psychometricians." This information is inaccurate in that our currently employed psychometrician, [EMPLOYEE B], has been with the TDOE Office of Assessment in this position for the past seventeen years.

Thank you again for the opportunity to respond to the draft report. If there are questions, or if more information is needed, please contact Dan Long (dan.long@tn.gov) or Chris Steppe (chris.steppe@tn.gov).

Sincerely,

/s/

Timothy K. Webb, Ed.D.
Commissioner

cc: Dan Long, Executive Director, Assessment and Evaluation
Chris Steppe, Director of Internal Audit