



UNITED STATES DEPARTMENT OF EDUCATION
OFFICE OF INSPECTOR GENERAL

AUDIT SERVICES
Chicago/Kansas City Audit Region

Control Number
ED-OIG/A07M0007

May 20, 2013

Michael P. Flanagan
Superintendent of Public Instruction
Michigan Department of Education
608 West Allegan Street
P.O. Box 30008
Lansing, MI 48909

Dear Mr. Flanagan:

This **Final Audit Report**, "Michigan Department of Education's System of Internal Control Over Statewide Test Results," presents the results of our assessment of selected aspects of the systems of internal control over statewide test results designed and implemented by the Michigan Department of Education (Michigan DOE) and three local educational agencies (LEAs). Our work in the State of Michigan (Michigan) was part of a nationwide audit of the systems of internal control over statewide test results put in place by the U.S. Department of Education (Department) and five State educational agencies (SEAs). The purpose of the nationwide audit is to determine whether the Department and the five SEAs have systems of internal control that prevent, detect, and require corrective action if they find indicators of inaccurate, unreliable, or incomplete statewide test results. The Office of Inspector General will issue an audit report that presents the results of the nationwide audit to the Department.

In Michigan, we performed our audit work at the Michigan DOE, Cesar Chavez Academy (Cesar Chavez), Detroit Public Schools (Detroit), and the School District of the City of Inkster (Inkster). Our audit covered statewide test results for school years 2007–2008 through 2009–2010. The purpose of this report is to separately address internal control weaknesses so that the Department and Michigan DOE can take appropriate corrective action before the next round of tests.

Michigan DOE and Detroit could improve their systems of internal control designed to prevent, detect, and require corrective action if they find indicators of inaccurate, unreliable, or incomplete statewide test results. Michigan DOE could improve its system of internal control by (1) placing schools that it identifies as high-risk for possible violations of test administration procedures on the next year's targeted monitoring list, (2) using test results and erasure analyses to identify schools with possible test administration irregularities, and (3) ensuring the timeliness of the missing nonscorable test materials reports that its contractors provide to schools and ensuring that its contracts are amended to include specific requirements for reporting missing nonscorable test materials (see Finding No. 1). Detroit could improve its system of internal

control by (1) adequately securing test materials, (2) retaining monitoring visit reports, and (3) emphasizing to schools that Michigan DOE and Detroit require schools to test students in a continuous session and report any deviations from required test administration procedures (see Finding No. 2).

BACKGROUND

The Elementary and Secondary Education Act of 1965, as amended, (ESEA) requires States to establish a set of high-quality, yearly student academic tests. The tests must measure the proficiency of students in mathematics, reading or language arts, and science. States must establish a single minimum percentage of students who are required to meet or exceed the proficient level on these tests. States use these tests to determine the yearly performance of the SEA, each LEA, and each school in the State. Section 1111(b)(3)(C)(iii) requires that the tests be valid, reliable, and consistent with relevant, nationally recognized professional and technical standards.

The Michigan DOE developed five tests that are used to measure SEA, LEA, and school performance:

1. The Michigan Educational Assessment Program (MEAP) is given to students in grades 3 through 9.
2. The MEAP-Access is given to students with disabilities.
3. The MI-Access is given to students with cognitive impairments.
4. The English Language Proficiency Assessment (ELPA) is given to students who are eligible for English language learner services.
5. The Michigan Merit Examination (MME) is given to students in grade 11 and eligible students in grade 12.

Students took the MEAP, the MEAP-Access, and the grades 3 through 9 MI-Access in the fall and the MME, the ELPA, and the grade 11 MI-Access in the spring.

Michigan DOE developed all the tests but worked with contractors to provide them. Contractors printed the test materials and distributed them to schools. Schools administered the tests, and then the LEA or schools returned completed tests and unused test materials to the contractors. The contractors scored the tests and provided the results to Michigan DOE, which made the results available to LEAs and schools via a secure Web site. Since school year 2008–2009,¹ the contractors also have provided Michigan DOE with erasure analyses that identify populations of

¹ In Michigan, classes run from September to June.

students or individual students with excessive erasures. Before school year 2011–2012, Michigan DOE considered a wrong-to-right erasure count that exceeds the State average by more than four standard deviations to be excessive.

Michigan DOE worked with the State Budget Office, Center for Educational Performance and Information, to upload the test results to *EDFacts*. *EDFacts* is a Department initiative to make performance data available for policy, management, and budget decisions for all K–12 educational programs. *EDFacts* centralizes K–12 performance data supplied by SEAs with other data, such as financial grant information, within the Department to enable better analysis and use in policy development, planning, and management. *EDFacts* includes data on student proficiency on statewide tests, participation rates on tests, and graduation rates at the SEA, LEA, and school levels.

As part of our LEA and school selection methodology, we used the *EDFacts* data to identify 663 LEAs in Michigan that had at least 1 school with total enrollment of more than 200 students during school years 2007–2008, 2008–2009, and 2009–2010. These 663 LEAs had 1,467,969 students attending a total of 2,769 schools that each had more than 200 students. Next, we calculated for each of the 2,769 schools a risk score for each grade tested in the subjects of math and reading. We calculated a risk score to determine how anomalous each increase or decrease in proficiency for a grade or subject was from one year to the next in relation to the change in proficiency for that grade or subject across the State.²

A maximum risk score was the highest risk score that a grade had for any of its subjects for the years that we reviewed. We considered a risk score greater than 10 to be anomalous. In a normal distribution, a risk score of 10 equates to a 1 in 10,000 chance of occurrence. We then selected the only three LEAs and five schools in Michigan that had multiple grades with maximum risk scores greater than 10. Because we used nonstatistical sampling procedures to select LEAs, our judgmental sample results were not projectable to all LEAs in Michigan.

Tables 1 and 2 show the number of schools and students for the LEAs and schools that we selected for review.

Table 1. Schools and Students in the LEAs Reviewed

LEA	Number of Schools	Number of Students
Cesar Chavez	3	1,601
Detroit	136	81,151
Inkster	3	2,559
Totals	142	85,311

² We used the risk score to select LEAs and schools for internal control reviews, not to determine whether cheating occurred at a particular LEA or school.

Table 2. Students in the Schools Reviewed

LEA	School Name	Number of Students
Cesar Chavez	Cesar Chavez Middle School	551
Detroit	John R. King Academic and Performing Arts Academy	867
Detroit	William Beckham Academy	690
Inkster	Baylor Woodson Elementary School	574
Inkster	Blanchette Middle School	567
Total		3,249

AUDIT RESULTS

Although Michigan DOE performed some internal control activities related to administering statewide tests, it could improve its systems of internal control. Michigan DOE provided LEA and school personnel annual training that explained proper test administration procedures. Michigan DOE also provided an “Assessment Integrity Guide” that contained handouts intended to assist LEAs with preparing and training their test administrators. Additionally, Michigan DOE’s Office of Standards and Assessment (OSA) conducted onsite monitoring of schools during test administration. However, Michigan DOE could improve its system of internal control over preventing, detecting, and taking corrective actions if it finds indicators of inaccurate, unreliable, or incomplete statewide test results. Specifically, Michigan DOE could strengthen its risk assessment and monitoring processes by—

- placing schools on its targeted monitoring list the year after it identifies them as being at a high risk for possible violations of test administration procedures and sharing the results of its onsite monitoring visits, if appropriate, with LEA and school officials;
- using reviews of test results and erasure analyses to identify schools with possible test administration irregularities; and
- ensuring the timeliness of the missing nonscorable test materials reports that its contractors provide to schools and ensuring that its contracts are amended to include specific requirements for reporting missing nonscorable test materials.

We also identified weaknesses in Detroit’s system of internal control over preventing, detecting, and taking corrective actions if it finds indicators of inaccurate, unreliable, or incomplete statewide test results that could call into question the validity of the test results. Detroit could improve its system of internal control by adequately securing test materials, retaining monitoring visit reports, and emphasizing to its schools that Michigan DOE and Detroit require schools to test students in a continuous session and report any deviations from required test administration procedures.

Though we identified internal control weaknesses in Detroit, we found that Cesar Chavez, Detroit, and Inkster generally were following the guidance that Michigan DOE provided to them. As Michigan DOE required, all three LEAs assigned people to perform the roles and

responsibilities of district test coordinator and building test coordinator. All three LEAs also provided yearly training to test administrators using Michigan DOE's guidance and required test administrators to sign the security compliance forms in the back of the test administration manuals.

In response to the draft of this report, Michigan DOE stated that it reviewed the findings and recommendations included in the report and discussed them with Detroit. Both Michigan DOE and Detroit agreed with the findings and recommendations. In addition, Michigan DOE described the corrective actions that have already been initiated and those it plans to initiate to address our recommendations. See Attachment 2 for the full text of Michigan DOE's comments.

FINDING NO. 1 – Michigan DOE Could Strengthen Its Risk Assessment and Monitoring Processes

Michigan DOE established a risk assessment process, and it followed a protocol to identify schools for targeted monitoring. It also developed procedures for monitoring its contractors, such as holding weekly or monthly meetings with contractors and performing data quality checks of its contractors' work. However, Michigan DOE could strengthen its risk assessment and monitoring processes by (1) placing schools that it identifies as high risk for possible violations of test administration procedures on the next year's targeted monitoring list and sharing the results of its onsite monitoring visits, if appropriate, with LEA and school officials; (2) using the student test results and erasure analyses that its contractor provides to identify schools with possible test administration irregularities; and (3) ensuring the timeliness of the missing nonscorable test materials reports that its contractors provide to schools, ensuring that its contracts are amended to include specific requirements for reporting missing nonscorable test materials, and following up with Detroit to determine what happened to Detroit's missing nonscorable test materials.

Michigan DOE Did Not Always Monitor Schools That It Identified as High-Risk Schools

Since school year 2007–2008, Michigan DOE's OSA has conducted onsite monitoring visits to schools during the MEAP test and MME. OSA employees recorded their observations on checklists that Michigan DOE periodically updated. OSA employees randomly chose some schools for visits and placed some schools with past test administration irregularities or invalidated test scores on a targeted monitoring list. However, Michigan DOE did not routinely place schools that it identified as high risk for possible test administration irregularities or test security procedures violations on the next year's targeted monitoring list.

Michigan DOE started electronically tracking reports of test administration irregularities in 2009. LEAs and schools reported these test administration irregularities. Irregularities reported by LEAs and schools included copying one year's test and using it to prepare students for future tests, helping students with questions, and giving students the wrong accommodations during testing. Michigan DOE's Assessment Administration and Reporting unit reviewed and analyzed all test administration irregularity reports and recommended an appropriate course of action.

In 2010, Michigan DOE added a severity coding system to its review of the test administration irregularity reports. Each report is assigned one of the following four severity codes:

- 0 = no misadministration occurred,
- 1 = unintentional misadministration with no threat to test validity,
- 2 = unintentional misadministration with compromised test validity, and
- 3 = intentional violation of administration standards with compromised validity.

The Assessment Administration and Reporting unit forwarded all test administration irregularities with a severity code of 3 to the ethics coordinator and the assessment administration and reporting manager, who reviewed the irregularity and determined a course of action. According to Michigan DOE, the types of action that the ethics coordinator and the assessment administration and reporting manager might take include (1) allowing the administration of an emergency test, (2) invalidating test scores, or (3) providing information about the irregularity to Michigan DOE investigators for further inquiry. In 2010, the Assessment Administration and Reporting unit assigned a severity code of 3 to test administration irregularities reported by eight LEAs. Of the eight LEAs, Michigan DOE invalidated the scores for three, approved the administration of emergency tests for three, referred one for further investigation by the Office of Educational Assessment and Accountability, and determined no action was required for one. However, Michigan DOE did not conduct onsite monitoring visits at any of the eight LEAs in 2011.

Michigan DOE agreed that it should have placed some schools with reported test administration irregularities in one year on the targeted monitoring list for the next year. However, the tracking system in place at the time of our audit did not record decisions on whether future monitoring was needed. Michigan DOE informed us that it plans to improve its electronic tracking practices by storing, in a searchable database, all onsite monitoring reports and determinations on whether additional action is needed.

Additionally, Michigan DOE did not share the results of any onsite monitoring visits with LEA officials, according to the Cesar Chavez, Detroit, and Inkster officials that we interviewed. Therefore, LEA officials could not immediately remedy potential test administration irregularities. Provided it shares monitoring visit results, if appropriate, with the LEAs and ensures that the LEAs take corrective action to address any test administration irregularities, improved monitoring procedures will help Michigan DOE ensure that LEAs and schools are properly administering statewide tests.

Michigan DOE Has Not Effectively Used Contractor-Provided Reviews of Test Results and Forensic Analyses to Identify Schools With Possible Test Administration Irregularities

Since school year 2008–2009, Michigan DOE’s statewide test contracts have contained provisions for contractors to provide Michigan DOE with student test results and erasure analyses, costing about \$20,000 annually, that identify populations or students with excessive erasures. Before school year 2011–2012, Michigan DOE considered a wrong-to-right erasure count that exceeded the State average by more than four standard deviations to be excessive. For the grades and subjects covered by this audit, this equates to an average of five wrong-to-right erasures per test. Although Michigan DOE’s Office of Psychometrics, Accountability, Research, and Evaluation received the student test results and erasure analyses, it did not (1) use the student test results to identify or follow up on unexpected year-to-year changes or anomalies in test scores at the LEA or school level or (2) follow up with LEAs or schools regarding populations or students identified with excessive erasures.

Michigan DOE provided us with school year 2008–2009 and school year 2009–2010 erasure data for the five schools that we visited. We compared the wrong-to-right erasure counts for each student at the subject, grade, and school levels to the statewide average. Using a statistical analysis,³ we determined the probability that a student would have a specific number of wrong-to-right erasures for each test. We considered a number of wrong-to-right erasures with a probability of occurrence of 1 in 10,000, which equates to an average of 13 wrong-to-right erasures per test, to be anomalous. Two of the five schools had one class with at least 20 percent of its students having an anomalous number of wrong-to-right erasures.

We then compared the erasure count results with our analysis of changes in student proficiency from one year to the next. Two schools with anomalous wrong-to-right erasure counts on tests also had anomalous changes in student proficiency for those tests. There are a variety of possible explanations for anomalous wrong-to-right erasure counts and changes in proficiency, and sanctions should not be based solely on data analysis. However, anomalous results on two different analyses at the same school demonstrate a level of risk that warrants follow-up. Michigan DOE could have used the erasure data along with other analyses, such as an analysis of changes in student proficiency or its review of irregularity reports, to identify schools with possible test administration irregularities.

Michigan DOE informed us that it has assigned the manager and chief psychometrician of the Bureau of Assessment and Accountability, Measurement Research and Psychometrics unit, the task of heading up forensic data analysis. In July 2012, the Bureau of Assessment and Accountability, Technical Advisory Committee, endorsed the methodology that Michigan DOE used to analyze the 2011–2012 MEAP and MME erasures. Michigan DOE informed us that it will use the analysis to direct communication to LEA leadership regarding erasure anomalies. According to Michigan DOE, this will mark the first time that it initiates a review based solely on erasure data. Michigan DOE also stated that it plans to establish permanent procedures for erasure analysis reviews in subsequent years, subject to input from the Technical Advisory Committee and its executive leadership.

Michigan DOE Did Not Ensure That Its Contractor Provided Timely Reports on Missing Nonscorable Test Materials to Schools

After schools administer the tests, they must return to the contractor all test materials, including answer sheets, test booklets, audio CDs, and Braille booklets. Schools packaged and identified returned test materials as either scorable (answer documents) or nonscorable (for example, test booklets). If necessary, the contractor then provided the schools with reports on missing test materials. These reports identified all scorable and nonscorable test materials that the contractor distributed to the schools but that the schools never returned to the contractor. Detroit administered the MEAP and MEAP-Access in October 2011. However, the contractor did not provide Detroit with a report that identified missing nonscorable test materials until May 2012, about 7 months after Detroit administered the MEAP and MEAP-Access.

The missing nonscorable test materials report included lists of materials shipped directly to the schools and lists of overage materials shipped directly to the LEA office. Some Detroit schools did not return one item; other schools did not return hundreds of items. Overall, 69 schools did

³ For details on the statistical analysis used, see the Scope and Methodology section of this report.

not return at least 1 nonscorable MEAP test item, and 10 schools did not return at least 1 nonscorable MEAP-Access test item. Because Detroit did not timely return these test materials, those in possession of the test materials had an extended amount of time to review and evaluate them, potentially compromising the integrity of future test results.

Michigan DOE's contract for MEAP administration states,

Once a school or district has indicated they have shipped all their materials then the Administration Contractor [Contractor(s) awarded this contract] will proactively review return and shipped materials logs to identify any missing materials then contact the school/district if any materials are determined to be missing. No later than one (1) week following the end of the assessment window the Administration Contractor shall also identify and contact all schools or districts that have not returned materials.

According to Michigan DOE, the 1-week requirement for the contractor to identify and contact all schools that have not returned test materials was intended to apply only to LEAs that did not return all scorable test materials. According to Michigan DOE, the contractor must process scorable test materials first to expedite the production of test results. Michigan DOE stated that the 1-week requirement cannot apply to nonscorable test materials because the process to inventory the nonscorable test materials is time-consuming. The MEAP has more than 5 million items that the contractor must scan individually. If the school has missing nonscorable test materials, the contractor asks the school to locate them. If the school cannot locate the nonscorable test materials, then the school must provide a written explanation. Based on the explanation that the school provides to the contractor, Michigan DOE might choose to take further action, which could include targeted monitoring of or investigating the school.

Federal Law, Regulations, and Guidance Require Monitoring for Compliance With Federal Requirements

Section 1111(b)(3) of the ESEA requires States to implement a set of yearly student academic tests that are valid, reliable, and consistent with relevant, nationally recognized professional and technical standards. According to "Key Policy Letters from the Secretary or Deputy Secretary," June 24, 2011, States are urged to make test security a high priority. States should review and, if necessary, strengthen efforts to protect student achievement and accountability data, ensure the quality of those data, and enforce test security. According to 34 C.F.R. § 80.40(a), grantees are responsible for monitoring grant- and subgrant-supported activities to ensure compliance with applicable Federal requirements and the achievement of performance goals.

Recommendations

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

- 1.1 Strengthen its risk assessment and monitoring processes by (a) placing schools that it identifies as high risk for possible test administration irregularities and test security violations on the next year's targeted monitoring list, (b) using the test results and erasure analyses that its contractor provides to identify schools at risk for test administration irregularities, and (c) ensuring that its contractor provides any required missing nonscorable test materials reports to schools in sufficient time to protect the integrity of future test results.

- 1.2 Follow up with the two schools with anomalous wrong-to-right erasures and anomalous changes in student proficiency to determine whether test administration irregularities occurred.
- 1.3 Share the results of test monitoring visits, if appropriate, with LEA and school officials and ensure that the officials take corrective action to address any test administration irregularities.
- 1.4 Ensure that its assessment contracts are amended to include specific requirements for reporting missing nonscorable test materials.
- 1.5 Follow up with Detroit to determine what happened to the unreturned nonscorable test materials.

Michigan DOE Comments

Michigan DOE agreed with the finding and recommendations and stated that it already implemented, or has plans to implement for the fall 2013 testing, corrective actions for all issues identified in the finding. Additionally, Michigan DOE stated that it has changed its approach to conducting erasure analysis and no longer uses a threshold of four standard deviations above the State average on wrong-to-right erasures to flag anomalous results. As described in its response, Michigan DOE now identifies students who have erasures of any type that exceed two standard deviations above the State average. From this group, it conducts further analyses and has a protocol in place to follow up on possible issues. According to Michigan DOE, using this new approach resulted in the identification of 10 schools for further follow-up after the fall 2011 administration of the MEAP. Michigan DOE further stated that many of the schools that the OIG identified with anomalous results were also identified using Michigan DOE's new process.

OIG Response

We consider Michigan DOE's comments to be responsive to our finding and recommendations. Additionally, in response to Michigan DOE's comments, we clarified in the finding that Michigan DOE used a threshold of four standard deviations above the State average to determine whether the number of wrong-to-right erasures was excessive before school year 2011–2012. We did not make any other substantive changes to this finding.

FINDING NO. 2 – Detroit Could Strengthen Its System of Internal Control Over Test Material Security, Recordkeeping, and Testing Administration

Detroit could strengthen its system of internal control over preventing, detecting, and correcting inaccurate, unreliable, or incomplete statewide test results by (1) adequately securing test materials, (2) retaining records of test monitoring reviews, and (3) emphasizing to its schools that all students are to be tested in a continuous session.

Detroit's Building Security Allowed for Unauthorized Access to Test Materials

During our site visit (May 30, 2012, through June 4, 2012), we noticed that unauthorized personnel had access to the area where Detroit stored test materials. Detroit employees were able to access the area simply by swiping their identification cards. On June 8, 2012, Detroit told us that it had disabled the identification card access on the level where it stored test materials,

and the area is now accessible only by key. However, later that same day (June 8, 2012), we noticed that the door did not lock automatically after it was unlocked with the key. Instead, somebody had to physically lock the door with the key to prevent unauthorized personnel from entering the secure area. After we brought this concern to its attention, Detroit stated that it would change the locks so that the door locks automatically after someone unlocks the door with the key.

Section 1111(b)(3) of the ESEA requires States to implement a set of yearly student academic tests that are valid, reliable, and consistent with relevant, nationally recognized professional and technical standards. “Key Policy Letters from the Education Secretary or Deputy Secretary,” dated June 24, 2011, stresses that State and local officials share responsibility for preventing threats to data quality and security breaches. State and local officials should review and, if necessary, strengthen their efforts to protect test and accountability data, ensure the quality of those data, and enforce test security.

Improperly or inadequately securing test materials can lead to missing test materials like the ones discussed in Finding No. 1 and increases the risk of inaccurate and unreliable test results.

Detroit Did Not Retain Records of Its Onsite Monitoring Visits

According to a Program Associate I with the Detroit Office of Research, Evaluation, Assessment, and Accountability, Detroit conducted onsite monitoring visits during the administration of tests at selected schools. The purpose of these visits was to monitor the selected schools’ test administration procedures. Reviewers completed monitoring checklists during each visit. However, when we asked for the checklists that were completed for our audit period, the Program Associate I informed us that the previous director of the Office of Research, Evaluation, Assessment, and Accountability either threw away the monitoring checklists from the prior years or took them with her when she left.

The same Program Associate I informed us that onsite monitoring visits were made to 19 schools during the administration of the 2011 MEAP; however, the Program Associate I could not find the completed monitoring checklists for 18 of the 19 schools. Without the monitoring checklists, Detroit was not able to verify whether the onsite monitoring visits were completed or act on any potential internal control weaknesses that the reviewers identified.

According to 34 C.F.R. § 80.42, grantees must retain all financial and programmatic records, supporting documents, statistical records, and other records reasonably considered as pertinent to program regulations for 3 years.

One Detroit School Did Not Test All Students in a Continuous Session

One Detroit school that we visited did not test students in a continuous session. The school allowed students who could not complete the scheduled section of the test before lunch to leave the test room to go to lunch at the same time as other students who had already completed the test. After lunch, the students requiring more time returned to the test room and completed their tests.

Schools cannot ensure the validity of test results for students who are not tested in continuous sessions. According to “Grade 4 MEAP Michigan Educational Assessment Program Test Administrator Manual Fall 2011,”

All parts of the MEAP test are untimed and student-paced. Students must be given as much time as needed during the same continuous session on the test date to complete each part of a test. . . . The test session may end when all students are finished. If only a few students need more time to finish, their test materials may be collected and they may be escorted immediately to a location where they may complete their tests. . . . Students who leave a room for an extended length of time (i.e., lunch hour, recess, etc.) should not be allowed to resume testing.

The manual also states that a student’s test results will be invalid if the school allows a student to take a lunch break.

Detroit’s “MEAP FALL 2011 TEST MEMO” included the same requirement for continuous testing that Michigan DOE included in the “Grade 4 MEAP Michigan Educational Assessment Program Test Administrator Manual Fall 2011.” The memo also states that allowing students to leave an assessment session for lunch is an example of an unacceptable activity.

The principal told us that students who had not finished the test were separated in the lunchroom from the students who had finished the test. Although students might have been separated in the lunchroom, the school did not follow the procedures required by Michigan DOE and Detroit. According to Michigan DOE’s “Assessment Integrity Guide,” when deviations from administrative procedures occur, the school should report the deviations to the district coordinator. Detroit’s “MEAP FALL 2011 TEST MEMO” also states that schools should immediately report unacceptable activities to Detroit. In this case, the principal did not notify Detroit of the deviation.

Recommendations

We recommend that the Assistant Secretary for Elementary and Secondary Education require Michigan DOE to ensure that Detroit—

- 2.1 Implements adequate security over test materials, including limiting access to the area where test materials are stored.
- 2.2 Retains all records of its onsite monitoring visits for at least 3 years and tracks and follows up on any test administration irregularities.
- 2.3 Emphasizes to schools that Michigan DOE and Detroit require schools to test students in a continuous session and report any deviations from required test administration procedures.

Michigan DOE and Detroit Comments

Michigan DOE and Detroit agreed with the finding and recommendations. Detroit stated that it installed a lock on the door to the room where it stores test materials, will submit electronic monitoring forms to Michigan DOE and retain the forms at Detroit, and will emphasize to schools that it requires schools to test students in a continuous session.

SCOPE AND METHODOLOGY

To achieve our objective, we—

1. Reviewed and gained an understanding of Federal laws, regulations, and guidance applicable to the audit objective.
2. Reviewed and considered the results of prior (1) Single Audits for Michigan DOE and the three LEAs and (2) Department program monitoring reports to identify areas of potential internal control weaknesses related to our audit objective.
3. Reviewed written policies and procedures, contracts for scoring tests, and assessment records at Michigan DOE, three LEAs, and five schools.
4. Interviewed officials at Michigan DOE, three LEAs, and five schools.
5. Gained an understanding and assessed the adequacy of Michigan DOE's and the three LEAs' systems of internal control over statewide test results in the following areas: data collection, monitoring procedures, and guidance.

We also analyzed erasure data for the five schools that we visited as part of this audit. We analyzed the data and calculated the probability of seeing the observed wrong-to-right erasure count for each student. Though some States might calculate the probabilities of wrong-to-right erasure counts using a normal distribution, we found that the Michigan erasure count data did not fit a normal distribution. Therefore, we used a distribution that better fit the observed data to support our stated probabilities. We calculated the probabilities using a generalized Poisson distribution based on the average statewide wrong-to-right erasure counts. To detect anomalous erasure counts, we flagged any wrong-to-right erasure count with a probability of occurrence of 1 in 10,000, an average of 13 wrong-to-right erasures per test.

Sampling Methodology

We judgmentally selected Cesar Chavez, Detroit, and Inkster from a universe of 663 LEAs. The 663 LEAs in our universe consisted of all LEAs in Michigan that had at least 1 school with more than 200 students during school years 2007–2008, 2008–2009, and 2009–2010. Next, we calculated for each school a risk score for each grade tested in the subjects of math and reading. We calculated a risk score to determine how anomalous each increase or decrease in proficiency for a grade or subject was from one year to the next in relation to the change for that grade or subject across the State. A maximum risk score was the highest risk score that a grade had for any of its subjects for the years we reviewed. We considered a maximum risk score greater than 10 to be anomalous. In a normal distribution, a risk score of 10 equates to a 1 in 10,000 chance of occurrence. We then selected the three LEAs and five schools in Michigan that had multiple grades with maximum risk scores greater than 10.⁴

⁴ We used the risk score to select LEAs and schools for internal control reviews, not to determine whether cheating occurred at a particular LEA or school.

Data Reliability

To achieve our objectives, we relied on data from *EDFacts*. *EDFacts* includes data fields for student proficiency on statewide tests, participation rates, and graduation rates at the SEA, LEA, and school levels. We used this data to select LEAs and schools to visit as part of this audit. We looked for patterns in the changes in student proficiency and used enrollment data to limit the universe of schools to those with more than 200 students during school years 2007–2008, 2008–2009, and 2009–2010.

To determine whether the *EDFacts* data were accurate and complete, we reconciled proficiency scores for selected grades and subjects at selected schools that we calculated using data that Michigan DOE provided us with the scores that we obtained from *EDFacts*. We did not find any discrepancies between the scores recorded in *EDFacts* and the scores that we calculated using Michigan DOE's data. Therefore, we determined that the data from *EDFacts* were sufficiently reliable for our intended use.

We also relied on erasure data that Michigan DOE provided. The data included fields identifying the LEA, school, grade, subject, erasure count, wrong-to-right erasure count, and standard deviation from the statewide average. The data also included the statewide average number of wrong-to-right erasures calculated for grades 3 through 8 in the subjects of math and reading. To determine whether the erasure data that Michigan DOE provided were reliable, we gained an understanding of Michigan DOE's processes for reviewing its contractors' scoring procedures. We also performed logic tests on the data. We looked for missing data and the relationship of one data element to another. Based on our understanding of Michigan DOE's processes and the results of our logic tests, we concluded that the data were sufficiently reliable for our intended use.

We conducted this audit at Michigan DOE's offices in Lansing, Michigan; Cesar Chavez's offices in Detroit, Michigan; Detroit's offices in Detroit, Michigan; Inkster's offices in Inkster, Michigan; and at our offices from March through December 2012. We discussed the results of our audit with Michigan DOE officials on September 5, 2012, and January 8, 2013. We provided Michigan DOE officials with a draft of this report on March 15, 2013.

We conducted this performance audit in accordance with the generally accepted government auditing standards (July 2007 revision). 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 on our audit objective. We believe that the evidence obtained provides a reasonable basis for our findings and conclusions on our audit objective.

ADMINISTRATIVE MATTERS

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 officials.

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 official, who will consider them before taking final action on this audit:

Deborah S. Delisle
Assistant Secretary
Office of Elementary and Secondary Education
400 Maryland Avenue, SW
Washington, D.C. 20202

It is the policy of the Department to expedite the resolution of audits by initiating timely action on the findings and recommendations contained therein. Therefore, receipt of any additional 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,

/s/

Gary D. Whitman
Regional Inspector General for Audit

Acronyms, Abbreviations, and Short Forms Used in this Report

Cesar Chavez	Cesar Chavez Academy
C.F.R.	Code of Federal Regulations
Department	U.S. Department of Education
Detroit	Detroit Public Schools
ELPA	English Language Proficiency Assessment
ESEA	Elementary and Secondary Education Act of 1965, as Amended
Inkster	School District of the City of Inkster
LEA	Local Educational Agency
MEAP	Michigan Educational Assessment Program
Michigan	State of Michigan
Michigan DOE	Michigan Department of Education
MME	Michigan Merit Examination
OSA	Michigan Department of Education, Office of Standards and Assessment
SEA	State Educational Agency



STATE OF MICHIGAN
DEPARTMENT OF EDUCATION
LANSING

RICK SNYDER
GOVERNOR

MICHAEL P. FLANAGAN
STATE SUPERINTENDENT

April 4, 2013

Gary D. Whitman
U.S. Department of Education
Office of Inspector General
500 West Madison Street
Suite 1414
Chicago, IL 60661

Mr. Whitman:

Thank you for the opportunity to respond to the "Draft Audit Report, Michigan Department of Education's System of Internal Control Over Statewide Results, Control Number ED-OIG/A07/M0007." We have reviewed the findings and recommendations in the report for the Michigan Department of Education (MDE) and have discussed these findings and recommendations for Detroit Public Schools (DPS) with Karen Ridgeway, Superintendent of Academics. MDE and DPS are in agreement with the findings and recommendations contained in this OIG report.

Furthermore, we agree with the corrective actions recommended in this report and in fact, are providing a description of the corrective actions that we have already initiated, as well as our plans to incorporate all of the remaining corrective actions suggested in this report.

We welcome this opportunity to closely review our Internal Control procedures and feedback you provided us to strengthen and improve our system. The audit team reviewed our activities from the academic years 2007 through 2010. Many of the Office of Inspector General's (OIG) recommendations and suggested actions were already being planned for, or were in various stages of implementation in the three academic years following the auditor's period (2011-2013).

STATE BOARD OF EDUCATION

JOHN C. AUSTIN – PRESIDENT • CASANDRA E. ULBRICH – VICE PRESIDENT
DANIEL VARNER – SECRETARY • RICHARD ZEILE – TREASURER MICHELLE
PECTEAU – NASBE DELEGATE • LUPE RAMOS-MONTIGNY KATHLEEN N.
STRAUS • EILEEN LAPPIN WEISER

608 WEST ALLEGAN STREET • P.O. BOX 30008 • LANSING, MICHIGAN 48909
www.michigan.gov/mde • (517) 373-3324

We describe in this plan our intention to:

- 1.1 Strengthen our risk assessment and monitoring processes by (a) identifying high risk schools for possible test administration irregularities and security violations on a targeted monitoring list, (b) using test results and erasure analysis and computer response anomalies provided by our contractor(s) to identify schools at risk for potential irregularities, and (c) ensure that our contractor provides MDE with a list of all schools that have nonscorable (hereafter in this document referred to as "secure") test materials that have not been returned and to aggressively pursue the return of those materials.
- 1.2 Since the audit, in the academic year 2011-12, we began follow-up on wrong-to-right erasure analysis on the Michigan Educational Assessment Program (MEAP) tests and have been expanding the practice to all other Bureau of Assessment & Accountability (BAA) statewide tests. The two schools identified in the OIG audit report were identified in our 2011-12 forensic analysis and we requested these school districts conduct a self-investigation and submit a written report. We have also begun efforts to conduct a state investigation at these sites. We continue to refine this process both for paper and pencil and are planning for computer-based administration procedures as well.
- 1.3 We believe the recommendation to share monitoring visit results with schools that are monitored is appropriate, and going forward, we will provide feedback to all schools we monitor and we are planning to change our Assessment Integrity Guidelines and BAA operational procedures to follow up with test administration irregularities or issues that result from monitoring or forensic analysis findings.
- 1.4 Beginning in the fall of 2013, we will improve our contractor and school administration requirements for reporting missing secure test materials.
- 1.5 We have had discussions and are following up with Detroit Public Schools (DPS) regarding unreturned secure test materials and have entered into planning with DPS to improve their and our response to this issue.

Detailed Operational Plans for each Recommendation.

Recommendation 1.1

Strengthen our risk assessment and monitoring processes by (a) identifying high risk schools for possible test administration irregularities and security violations on a targeted monitoring list, (b) using test results and erasure analysis and computer response anomalies provided by our contractor(s) to identify schools at risk for potential irregularities, and (c) ensure that our contractor provides MDE with a list of all schools that have outstanding secure missing secure test materials that have not been returned and to aggressively pursue the return of those materials.

(a) Identification and Monitoring of High Risk Schools.

We have already implemented an electronic means of tracking all irregularities and we have more than doubled the number of monitoring visits. The database will allow us to track all known irregularities and any schools on that list will be identified and flagged for monitoring. We have changed our monitoring procedures so that our contracted investigators conduct targeted monitoring of integrity risks and our test administration contractors target a random geographic list of schools for assessment administration quality monitoring. We are continuing to explore improvements in software to improve this data collection and monitoring system. We also have in place procedures to identify possible risks brought to our attention by providing a complaint (or allegation of misadministration) form on the main BAA website and by a toll-free number (877)560-8378. The first item on the routing call message for the toll-free number will remain with the instructions; "For issues related to inappropriate or unethical testing practices, please press 8." This will route the caller to our test administration coordinator who handles all testing irregularities. If that individual is not available, the system will invite the caller to leave a voice mail or the caller can route to the next person in our chain of command. This information is also collected in an electronic database so that all input sources lead to the same database with all detailed information.

(b) Forensic Analyses

The BAA recognizes and agrees with the findings of the auditors regarding forensic analysis and has already initiated these activities beginning in the academic year 2011. However, there is a factual finding regarding erasure analysis that is not reflective of our current and improved procedures. BAA no longer utilizes a threshold of four standard deviation units above the mean on wrong-to-right erasures to flag anomalous results. Using the four standard deviation procedure identifies very few issues due to the restrictive nature of that metric and is also confounded by interactions of sample size and type of test. We recognize that psychometric data forensics as they relate to testing irregularities is a new and evolving field of research which currently lacks widely agreed upon accepted practices. Just as multiple procedures exist for large-scale psychometrics (such as equating, scaling, standard setting, and alignment), multiple methods, models and approaches exist for forensic analysis that are emerging by the field and supported in research literature.

We continue to review the best current practices and emerging research. We are also partnering with other SEAs and university faculty with the goal of determining the best sets of analyses to use in Michigan. As the audit findings suggest, we have been and will continue consulting the Michigan Technical Advisory Committee (TAC) as we move from research, to proposal, to policy, to operations.

Starting with the fall 2011 administration, the BAA has used a multi-step procedure to flag schools for further investigation based on the results of the erasure analysis on the Michigan Educational Assessment Program (MEAP). For the fall 2012 administration, we expanded the erasure analysis process to include MEAP Access (MEAP alternative assessment), and in Spring 2013 we are expanding this analysis to the Michigan Merit Exam (MME).

The following erasure procedures have been approved by our TAC and have been found to be most efficacious. The analysis begins with individual student answer documents as the unit of analysis and the statewide mean and standard deviation (by grade, content and assessment) is calculated based on the total amount of erasures present as opposed to the wrong-to-right erasures only. Once those benchmarks have been established, students whose total erasure count exceeds 2 standard deviations are being flagged for possible risk. The rationale for the 2 standard deviation approach is to recognize that on many of these assessments, particularly the MEAP-Access and MME, erasures are not numerous and a higher threshold would be insensitive to our analysis. That subset of students for each grade/ content/ assessment is then further queried to determine which of those students' erasures are found to be at least 75% were wrong-to-right. The 75% threshold for students' wrong-to-right changes was based on the rationale that students who are not only excessively erasing but who also are very accurate in that erasure pattern need to be further evaluated to determine if this behavior is due to learned testing behavior (e.g. the student is told to take their time on the assessment and carefully go back over all of their answers in a session when completed) or could possibly be due to actions being carried out by students or others that might compromise the validity of interpretations from those specific students' test results. Of those students flagged, we also identify the school in which they were assessed - allowing us to aggregate to the school (and district if needed) level to determine if flagged students are spread in low numbers across multiple schools or if in fact, there is a preponderance of such anomalous results residing within the same building, and perhaps at the same grade and in the same classroom. Only the extreme cases (outliers) of the distribution of the number of flagged students within schools will be contacted regarding the results.

Our forensic analysis identified 10 schools following the Fall 2011 administration and BAA contacted them with a letter of inquiry. We received written responses from 8/10 schools. One school had changed names but is in the same location. The three schools from which we did not receive a written response are being assigned to our investigators. Coincidentally, many of the schools USED identified with anomalous results also showed up in our analysis. Therefore, we believe our method is robust enough to capture such anomalies and consistent with the USED practices.

Our future plans continue the aforementioned practice and we are working with our scoring and reporting contractors to move the delivery of the erasure analysis results to BAA Psychometrics as early as feasible in the assessment cycle. This should allow the BAA sufficient time to conduct the analysis and follow up with schools in a timely fashion. It is our goal to do this prior to the public release of reports so that if any corrective action is needed that would impact the integrity of the data, it can be carried out without impacting statewide results and reports. As suggested by USED, schools where incident reports are on file or contact has been made reporting inappropriate testing behavior, will also be subject to closer inspection of response patterns and erasure data, and from school monitoring observations.

We continue to research and develop a process for flagging anomalous score gains and we are working toward a procedure that will garner our TAC's approval. We have noted the Auditor General's methodological summary with great interest. While we do not believe the method employed by USED will be directly applicable for our data procedures, we do plan to conduct research using multiple methods of similar or equivalent analyses to determine which procedures will serve our purposes the best. We are in complete agreement with the audit findings that a conclusion of cheating or dishonesty should never be based solely on one analysis or approach. It is beneficial when reported allegations are substantiated by the data and by factual findings within a reasonable time frame.

In summary, it is our goal to minimize spurious claims of irregularities due to chance circumstances and we want to be very confident in our results before we escalate the potential irregularity and launch state investigations. Therefore, we are altering our procedures to include more detailed review of forensic analyses and incorporate identification, and various levels of inquiry (up to investigations and remediation). We are planning for the use of multiple indicators and will implement these approaches as we work through available options, obtain approval from our TAC, and put the procedures into operations as soon as possible.

(c) Aggressively pursuing the return of secure test materials from schools

As we put new test administration contracts in place for the 2013 school year, we will shorten the timeline for the reporting of unreturned secure materials. In addition, we are developing new methods to communicate these missing secure materials issues to the involved schools and exploring additional methods to improve the tracking and return of missing secure materials.

Recommendation 1.2

Since the audit, in the academic year 2011-12, we began follow-up on wrong-to-right erasure analysis on the Michigan Educational Assessment Program (MEAP) tests and have been expanding the practice to all other BAA statewide tests. The two schools identified in the OIG audit report were identified in our 2011-12 forensic analysis and we requested these school districts to conduct a self-investigation and submit a written report. We have also begun efforts to conduct a state investigation at these sites. We continue to refine this process both for paper and pencil and are planning for computer-based administration procedures as well.

We have developed and implemented a plan for erasure analysis, resulting inquiry, and follow up. Using the procedures described in 1.1 for erasure findings, schools that are flagged as at-risk will be sent a letter of notification indicating that we have found a significant erasure anomaly and ask the school district to conduct a self-investigation and respond to us. A risk assessment team comprised of BAA staff with direct responsibilities for assessment integrity meets and reviews the school's response. The team then chooses from among four options: 1) accept the findings of the school and their remediation of the problem, or MDE recommended remediation (such as invalidating the scores), 2) require more and specific analysis or additional fact finding, 3) place the school on the monitoring list for the next administration, or 4) recommend an state investigation by a BAA Independent Investigator. As a result of a review of the fall 2011 data, we had 10 flagged at-risk schools that were sent letters of inquiry, all 10 schools will be placed on the targeted monitoring list, and 3 schools will be investigated. We have moved up the fall 2012 testing erasure reports and we plan to conduct the resulting inquiries from this year's MEAP administration in May of this year.

Recommendation 1.3

We believe the recommendation to share monitoring visit results with schools that are monitored is appropriate and we will provide important feedback to all schools we monitor. We are planning to change our Assessment Integrity Guidelines and BAA operational procedures to follow-up on test administration irregularities or issues that result from monitoring or forensic analysis findings.

Beginning with the fall 2013 test administrations, BAA will modify the monitoring checklist to include a "needs improvement" category and monitors will be re-trained to ensure that, in addition to documenting acceptable and unacceptable practices, they will use the checklist to make suggestions for improvement.

In the past, these monitoring checklist forms have been filled out by hand and collected. Our current plan is for the reports to be entered into a web-based electronic form so that the forms can be easily emailed to the school building coordinators, and our collections will now be easier to access and search.

Recommendation 1.4

BAA has improved contractor and school administration requirements for reporting missing secure test materials. The following procedures will be in place for fall 2013 test administrations:

Schools are required to collect signed Security Compliance Agreement forms from all staff handling secure print materials in schools and districts. The signed Security Agreement Forms are required to be maintained on file by the school or district for a period of one year following the completion of the testing window. In addition to reviewing the documents when investigating testing irregularities, contractors visiting schools for the purpose of targeted or random test administration monitoring will review the forms as part of the monitoring checklist.

Scoring assessments is both time intensive and time sensitive. Immediately following the scoring of assessment materials the contractor now scans nonscorable secure materials and notifies schools of any missing materials. Prior to fall of 2013, this process had to be completed by the contractor by the end of May for fall assessments. Effective with the fall of 2013, the contractor will be required to complete the process by the end of February. BAA and the Contractor(s) are currently exploring ways to shorten this processing time for paper and pencil tests.

The fall of 2013 will be the last full paper and pencil administration of MEAP, MEAP-Access, and MI-Access. On-line piloting will be done with Social Studies and Science. Smaller numbers of schools using paper and pencil will make this task much easier to complete earlier in the process.

Second letters are sent to schools and districts who do not respond to the first letter of inquiry about missing materials. These schools are identified for possible monitoring during the next test cycle or can be referred to our independent investigators to explore this issue as a possible irregularity.

Online testing is expected to be available in the spring of 2015. Although paper and pencil tests will be an option, the handling of materials (both secure and non-secure) will significantly diminish as Michigan moves more fully to online assessment.

Recommendation 1.5

We have followed up with Detroit Public Schools (DPS) regarding unreturned secure test materials and have entered into planning with DPS to improve their and our response to this issue.

BAA communicates directly with DPS schools about missing materials. Although most testing materials across the state are sent directly to schools (and also returned by schools), DPS testing materials are sent to the district assessment office and distributed to the schools by DPS personnel. This process is designed to increase accountability and provide a clear chain-of-custody for materials. We have received assurances from DPS that they will be in full cooperation with this procedure.

BAA has discussed the Draft Audit Report with DPS and has received the following communique from Karen P. Ridgeway, DPS Superintendent of Academics:

"This communication is in response to the "Draft Audit Report, Michigan Department of Education's System of Internal Control Over Statewide Test Results, Control Number ED-OIG/A07M0007", Finding No. 2 and recommendations 2.1, 2.2 and 2.3. pages 9-11 of 14.

Finding No. 2 - Detroit Could Strengthen Its System of Internal Control Over Test Material Security, Recordkeeping, and Test Administration

2.1ensure that Detroit-
Implements adequate security over test materials, including limiting access to the area where test materials are stored.

DPS Response:

The District concurs with the information. The Office of Research, Evaluation, Assessment and Accountability requested that the locks to the identified area where the statewide assessment materials are housed be changed. This request was granted within 24 hours of the request. The lock is now automatic and the area can now only be accessed by the use of a physical key. The door locks automatically when it is closed. Swipe cards can no longer be used to access the area and the distribution of a physical key is limited to assessment staff.

2.2ensure that Detroit-
Retains all records of its onsite monitoring visits for at least 3 years and track and follow up on any test administration irregularities.

DPS Response:

The District will digitize its test administration onsite monitoring form and agrees to submit these forms to the Office of Standards and Assessment, Michigan Department of Education upon completion of any onsite visit. This action allows for an electronic record to be maintained at both the District and the Department. The District understands that it is responsible for follow-up on any test administration irregularities. It should be noted that the Detroit Public Schools regularly practices self-reporting in any event of test misadministration or irregularity. The District also understands that any test irregularities identified in the submissions of the onsite monitoring forms to MDE are subject to comment and action by MDE.

2.3ensure that Detroit-

Emphasize to schools that Michigan DOE and Detroit require schools to test students in a continuous session and report any deviation from required test administration procedures.

DPS Response:

Please be advised that the Detroit Public Schools publishes a "TEST MEMO" for each state test administered. Each "TEST MEMO" strongly advises all schools that once a test session begins, students are to complete testing during that single session. The Office of Curriculum and Instruction will require each school to submit an electronic copy of the testing schedule for each state assessment. These schedules will be reviewed for compliance with the identified requirement. The Office of Research, Evaluation, Assessment and Accountability will have access to each testing schedule as they conduct onsite visits.

The District will highlight this finding in all of its test administration training sessions and at each principals' meeting prior to the start of state testing. Additionally, this item will be included on the electronic test administration onsite monitoring form and will be reviewed by Office of Research, Evaluation, Assessment and Accountability staff. The principal of the identified school will be counseled and a special training session will be held at the school prior to the administration of the next state assessment.

Thank you for allowing the Detroit Public Schools to comment on these findings and recommendations. Please let me know if additional information is required."

Mr. Gary Whitman
Page 10
April 4, 2013

In closing, we believe that the MDE has either already incorporated all of the recommendations from the Draft Audit Report or will have such recommendations in place for the fall 2013 assessments.

Thank you for your assistance in this audit. If you have any questions, please direct them to Dr. Joseph Martineau, Executive Director of the Bureau of Assessment and Accountability, martineauj@michigan.gov, (517)241-4710.

Sincerely,

/s/

Michael P. Flanagan
State Superintendent

cc: Vince Dean, Director, Office of Standards and Assessment, BAA
Dave Judd, Director, Office of Systems, Psychometrics and Measurement Research, BAA
Joseph Martineau, Executive Director, BAA
Karen Ridgeway, Superintendent of Academics, Detroit Public Schools, DPS
Sally Vaughn, Deputy Superintendent, Chief Academic Officer, MDE