

# CHAPTER 10: QUALITY CONTROL PROCEDURES

## Summary

The statewide testing program and its data play an important role in public education accountability. The Texas Education Agency verifies the accuracy of the work done and data produced by the testing contractor, Pearson Educational Measurement, through a comprehensive verification system using three fictitious school districts that simulate different demographic categories and test-scoring scenarios. In addition, PEM has its own internal quality control system to verify the accuracy of the reports it produces for Texas school districts.

## Reporting of Test Results

Individual student test scores are used for promotion, graduation, and remediation. In addition, the aggregated student performance results from the statewide testing program are a major component of the state and federal accountability systems that are used to rate individual public schools and school districts in Texas. The data are also used in education research and in the establishment of public policy. Therefore, it is essential that the tests are scored correctly and that the data are accurately reported to school districts. PEM is responsible for scoring the tests, aggregating the results, and printing and shipping the reports to the school districts. To ensure the accuracy of the reported test data, TEA uses a comprehensive quality control system (QCS) to verify the accuracy of the data and reports produced by PEM. The QCS is implemented for every TAKS (both paper and online), TAKS-I, SDAA II, TELPAS, and TAAS exit level administration, and PEM participates with TEA throughout the QCS process.

In addition to interacting with TEA through the QCS, PEM has its own internal quality control system for the reporting of test results. PEM's internal quality control starts with the utilization of a business process known as the Capability Maturity Model (CMM). CMM is a description of the stages through which software organizations evolve as they define, implement, measure, control, and improve their software processes. This model provides a guide for selecting process improvement strategies by facilitating the determination of current process capabilities and the identification of the issues most critical to software quality and process improvement. CMM requires the creation of documents that assist in the requirements definition, development, testing, and implementation of the software required to support each testing program. Examples of these documents include a customer requirements allocation document, a project schedule, functional specifications, a software development project plan, unit test plans, and verification and validation plans.

Once software requirements have been gathered, project schedules, project plans, functional specifications, and design documents are created. From these, unit test plans and system test plans can be determined. A unit test plan is a list of code-unit test cases that is executed and

recorded by the software developer. The purpose of the code-unit test process is to ensure that software is developed, maintained, documented, and verified to meet the project requirements for coding and unit testing. As such, the process provides the mechanisms that are necessary to implement the software requirements and design as well as provides code-units quality assurance prior to system test.

After all modules (units) have been tested within a system, the CMM process requires a system test. The system test ensures that all the units work together and that outputs from one module match up to the proper inputs for the next module in the system. It also uses expected results to ensure that all requirements have been met. It is important that the system test be performed by a group that is independent of the software development group. This process allows for independent verification and interpretation of the requirements. Once the independent testing group has completed the test and given its approval, the system is moved into production mode. It is ready to process the QCS documents and files supplied by TEA, as described in the following paragraphs.

TEA begins the QCS months in advance of a test date. For each test administration PEM and TEA prepare answer documents for hundreds of fictitious students, who are assigned to a campus in one of three fictitious districts. PEM grids these answer documents (marks the answer choices and student identification information) using detailed instructions provided by TEA. The answer documents represent real-world scenarios of the numerous correct and incorrect ways answer documents are completed by students and districts. These scenarios may include but are not limited to

- gridding an answer document that has all questions answered correctly in all subjects;
- gridding an answer document that has an incorrect Public Education Information Management System (PEIMS) identification number;
- gridding all possible demographic categories (for example, special education status, gifted/talented status);
- gridding two score codes instead of one for a subject area;
- changing the student name on a precoded answer document;
- marking a score code of "A" (absent) for a student who took the test;
- verifying each of the possible raw score points (for example, in mathematics); and
- returning answer documents to PEM for scoring with incomplete or missing student identification information and/or improper Class and/or Campus Identification Sheet.

PEM then processes, scores, and prepares reports for these fictitious students using answer keys, editing rules, and formats approved previously by TEA. TEA simultaneously processes the same student-level information and produces its own reports. When TEA receives Pearson Educational Measurement's reports for the fictitious students and districts, it compares PEM's reports with its own reports.

In addition to scores, calculations, and other numerical data printed on the reports, all text, formats and customized messages are also verified. This procedure also verifies that edits are made properly when the scanner encounters missing or invalid data. Reports are not sent to districts until all discrepancies in the comparative data for the fictitious districts are resolved and the reports generated by TEA and PEM match. In addition, the verification system allows TEA to monitor the distribution of all test materials, reports, and information letters.

The following are more specific details of the QCS process for report verification.

1. Prepare a test design for each test administration. This is a set of specific instructions to PEM for preparing answer documents for fictitious students.
  - Check the proposed answer document for the upcoming administration for any design changes that might affect the QCS process (for example, new or revised data fields).
  - Determine whether any new policies have been established since the last administration of the test that would affect how answer documents are edited or how scores are reported. Decide how these policies affect the QCS process and whether these new edits should be tested with additional fictitious students.
  - Using Microsoft Access, create a new database of fictitious students. A new test administration will have most of the same students as the previous administration of the same test but with additions or changes necessary to reflect new reporting policies and/or new conditions that should be tested.
  - Prepare a written test design consisting of coding and gridding instructions to PEM.
    - Produce a spreadsheet by exporting a file to Microsoft Excel from the appropriate Access database. The spreadsheet provides (1) the instructions needed to code the answer document for each fictitious student, (2) information about which answer documents to precode, and (3) a list of which answer documents are to be gridded by hand and which are to be gridded electronically.
    - Create a Microsoft Word document with instructions for grouping the students under Campus and Class Identification Sheets and additional special coding instructions such as overwriting precoded data and marking more than one bubble in a data field (double-gridding). The Word document is divided into several sections, each of which is a group of students that test similar aspects of the reporting process. The Word document also provides district, campus, and class summaries of the numbers of students.
    - Export a selected portion of the database as a text file. This provides PEM with the data and instructions to grid some answer documents electronically. There are hundreds of students in each database for each administration, and not all data on all answer documents for fictitious students are gridded by hand.

- Send the written test design and text file to PEM per an approved schedule of processing deadlines created for the particular test administration.
2. Receive scales from PEM.
    - PEM sends a table to TEA that shows the scale score corresponding to each achievable raw score point. If a test administration uses pre-equated scales, these true scales will be used to convert the raw scores to derived scores and assign a passing status (TAKS), achievement level (SDAA II), or proficiency rating (TELPAS) for each fictitious student. These tables are verified, approved, and incorporated into computer programs that produce the student and district/campus files and reports.
    - If a test is post-equated, an artificial scale is used initially for processing the fictitious students' answer documents. Because the QCS is a lengthy process, waiting for the true scale to be created before verifying the accuracy of the reporting system would compromise the delivery schedule of reports to districts. For most of the spring tests, there are only 1–3 days between scale approval and sign-off for QCS. Therefore, there is a thorough comparison of files and reports (see below) using artificial scales and an abbreviated comparison of reports with scores generated with the post-equated, or true, scales.
  3. Create a student-level data file. This file contains the data from the simulation of the processing of answer documents from the fictitious students.
    - Verify that "resolved" fields are correct in the database. The resolved fields simulate the changes that would be made in the PEM editing process if coding or bubbling errors are made on the answer document.
    - Export the data from the database as a text file and create a SAS dataset.
  4. Receive student-level file from PEM.
    - PEM sends a student-level file (text file) to TEA with student names, demographic data, and scores. This file is the result of a procedure which, using the instructions in the test design provided by TEA, simulates the bubbling of answer documents by districts, followed by processing, editing, and scoring of answer documents by PEM. These data are in the format of the Electronic Individual Student Record File, which is an optional report available to districts.
    - Create a SAS dataset from PEM's student-level text file.
  5. Compare PEM and TEA files.
    - For each student record compare each variable in the PEM student-level data set with the corresponding variable in the TEA student-level data set.
    - Investigate each mismatch, if any, and determine the source and reason for the discrepancy.
    - Make corrections, if necessary, in accordance with established policy and edit rules.

- Repeat the process by regenerating student-level files, comparing and resolving discrepancies until the files are identical.
6. Print reports.
    - PEM prints reports for the three fictitious districts (various types of both individual and aggregated reports) and sends them to TEA.
    - A corresponding TEA report is produced for each PEM report.
    - Reports by TEA and PEM are generated with independently produced computer programs.
  7. Verify reports.
    - Reports are compared to verify that they contain identical information.
    - Any discrepancies found are investigated and corrected.
  8. Approve reports.
    - When all the reports for the fictitious districts are verified to be free of error, TEA notifies PEM that reports can be shipped to school districts when PEM's quality assurance process is complete.

When all the standard and optional reports for the fictitious districts have been verified by TEA, the system is then available for the processing of the school districts' documents. Before the bulk of districts' documents is allowed to be processed, the first production run process (FPRP) is performed. A small representative sample of documents is readied for processing and is processed through the entire system. The FPRP documents record a sequence of defined activities used in the first run of live production data through PEM's Operations Department functions that are specified for a program/project. The FPRP serves as the transition point from the planning, development, and testing phases to the delivery phase, and from software development and testing to the production environment. The formal process allows all participating functional groups to

- formally accept the readiness of a system for full production;
- confirm receipt and comprehension of processing specifications; and
- confirm the receipt of required production materials for the project.

Once the FPRP is complete and PEM's Operations and Quality Assurance departments have approved the functions, PEM completes the scoring and reporting process for the districts. PEM also ships final reports to the addresses of the fictitious districts on the same schedule as shipments to the actual districts, and TEA monitors these shipments for timeliness and completeness.

# Ongoing Quality Control

## Content Validation

During the 2003–2004 school year, the Texas Education Agency and Pearson Educational Measurement added an additional quality control step in the item development process. Content validation review is now an established step in item quality control procedures. In 2005–2006 panels composed of university-level experts in the fields of English language arts, mathematics, science, and social studies were assembled in Austin to review the content of the high school level TAKS assessments before their administration. This quality control measure was instituted to ensure that each high school test was of the highest possible quality. The review was deemed necessary at the high school grades (9, 10, and 11) because of the advanced level of content covered on the assessments. In a series of one-day meetings scheduled for each content area, experts independently reviewed the test forms. After a thorough review of each test, committee members discussed all items, noting any issues that were of concern. If necessary, substitute items were chosen and reviewed. Committee member recommendations were then shared with management for resolution.

## Verification of Equating Activities

As another quality control step, PEM performs verifications of all equating activities. Each live TAKS test is calibrated and equated by four people: two PEM psychometricians, a TEA psychometrician, and an external psychometrician. An additional PEM psychometrician acts as a Quality Control Coordinator and reviews the four sets of equating results. Any discrepancies across the results are resolved prior to the generation of the final scoring tables. In addition, the Quality Control Coordinator compares the current year's post-equated results to those from previous years and conducts additional data quality and reasonableness checks. The SDAA II and RPTE test calibration and equating procedures are conducted independently and verified by two PEM psychometricians with results reviewed by an additional PEM psychometrician acting as a Quality Control Coordinator. Field-test items for all three testing programs, whether embedded or separately field-tested, are also independently calibrated and equated by two PEM psychometricians. Equating is discussed in detail in Chapter 16.