

WEST VIRGINIA

Quality Control Checklist for Processing, Scoring, and Reporting

1. Test keys and items to standards/objectives maps are accurate.

Check √	Stage/ Task Purpose	Task	Recommended Expertise *			Recommended Materials	
<input type="checkbox"/>	<u>During test form development</u> To ensure the match of test items and content standards/objectives	1A. During test form development, test items are selected based on the test blueprints. Each selected item aligns to the content standard and objectives. WVDE staff in each content area confirms the match of each test item to the content standard/objectives and confirms that the test matches the test blueprint and meets the content coverage requirement.	C	A	P	Content standards/objectives, item cards, test blueprints, assessment framework	
<input type="checkbox"/>	<u>During test form development</u> To ensure accuracy of information on the test maps	1B. Item cards and test maps are created to document the information for each item selected for final test production. The information includes each item's correct response, the maximum number of score points, the item's intended standard/objective, and other attributes (e.g., location, field test data which includes p-value and point bi-serial). WVDE staff (1) verifies the accuracy of PEID and all information on the test maps before they are sent to the vendor for test booklet production, (2) ensures that all items for each test form meets the statistical requirements, and (3) uses CTB's Itemwin to create equated test forms in each of the four content areas.	C	A	P	Item cards, test maps, PEID, test blueprints, item records, field test results, psychometric rules for test development	
<input type="checkbox"/>	<u>After test form development</u> To ensure accuracy of answer keys for multiple-choice items and accuracy of information for open-ended items	1C. To verify the accuracy of the answer key information for multiple-choice items on the test maps, WVDE staff takes each test for their respective content area and compares their answers to the answer key on the test maps and in PEID. They also verify PEID item-to-standard/objective information on the test maps and in PEID. To verify the accuracy of the information for open-ended items, WVDE staff reviews the items, develops scoring rubrics with CTB, and approves the maximum score point assigned to each open-ended item, the item-to-standard/objective classification for each item, and other item attributes.	C	A	P	Item cards, PEID, test maps, test items, scoring rubrics, handscoring rules	
<input type="checkbox"/>	<u>After scanning test booklets</u> To verify accuracy of answer keys and item-to-standard/objective codes in the data file that is used to produce reports	1D. Using a sample of actual student response data prior to the operational scoring process, WVDE staff verifies the accuracy of answer keys in the data file (PEID) provided by the vendor. Frequency distributions of item responses or point-biserial information are used to check multiple-choice answer keys; frequency distributions of score points for each item could be used to check students' performance on the open-ended items. Observations of unexpected results should be investigated and brought to resolution.			P	D	Sample student responses file, PEID, test maps

2. Student responses are completely and accurately scanned and captured.

Check √	Stage/ Task Purpose	Task	Recommended Expertise*			Recommended Materials	
<input type="checkbox"/>	<u>Before scanning test booklets</u> To ensure the vendor's scanning system works appropriately and accurately	2A. WVDE includes procedures in the Test Coordinators Manual that facilitate the accuracy in scoring. Prior to the scanning and scoring of actual response documents, WVDE requires the vendor to provide a "test deck" file and a copy of the "test deck" answer documents. The test deck procedures involve processing hand-bubbled test booklets with a variety of scenarios through the scanning process allowing simulated student test data to be processed and scored. WVDE staff compares the information on the test deck file with the bubbles on the simulated test booklets and reviews the reports provided by the vendor.		A	P	D	Test deck file, test deck reports, simulated test documents
<input type="checkbox"/>	<u>Before scanning test booklets</u> To ensure the vendor's imaging system works appropriately and accurately	2B. Prior to the scanning and scoring of actual response documents, WVDE requires the vendor to (1) demonstrate that student responses are captured completely and are readable on-screen and when printed, (2) specify procedures for re-scanning documents if images are determined to be unreadable during scoring, and (3) demonstrate that individual response images distributed for scoring are accurately linked to the correct student record.	A				Quality assurance report on the imaging system
<input type="checkbox"/>	<u>During the scanning process</u> To ensure ALL student responses are scanned and captured	2C. WVDE requires the vendor to provide information regarding the status of the test documents throughout the scoring and editing process. All test documents and school/student headers for all expected schools are received and scanned. Actual counts based on scanned records are compared with expected counts from the WVEIS student and school files. Any discrepancies are researched and resolved. WVDE staff reviews the documents provided by the vendor and assists with the investigation of any discrepancies.		A	P	D	Scanning and editing status reports
<input type="checkbox"/>	<u>During the scanning process</u> To ensure that student responses to multiple choice items are authentic student responses	2D. The vendor maintains installed technologies to conduct erasure analyses. Erasures of student answers can be identified electronically and logged. The frequency of erasures is monitored and an alert system notifies the vendor when the number erasures are statically abnormal. WVDE has in place policy practice to review and investigate the discrepancies of student responses.		A	P	D	Erasure analysis technologies
<input type="checkbox"/>	<u>After the image scanning process</u> To ensure accuracy of scanned images of responses for open-ended items	2E. WVDE requires the vendor to deliver (1) a set of answer documents produced from scanned images and (2) photocopies of the same set of answer documents. WVDE staff should verify the accuracy of information on the scanned images against the photocopies of the original documents and also ensure that the student response images are matched back with the proper student's record.		A	P	D	Imaging materials
<input type="checkbox"/>	<u>After the scanning process</u> To ensure accuracy of scanning data for multiple-choice items	2F. WVDE requires the vendor to deliver (1) a scanning file including scanned multiple-choice responses and (2) a copy of a set of answer documents for each test form. WVDE staff verifies the accuracy of information on the scanning file against the bubbles on the test booklets.		A	P	D	Scanning file, copies of answer documents

3. Student multiple-choice responses are accurately scored.

Check √	Stage/ Task Purpose	Task	Recommended Expertise*		Recommended Materials	
<input type="checkbox"/>	<u>After scanning and scoring multiple-choice (MC) items</u> To ensure accuracy of the scoring data for multiple-choice items	3A. Prior to the final processing of student test results, WVDE requires the vendor to provide the preliminary data file for quality control of the scoring and reporting processes. If the state file is extremely large, the vendor provides a sample student responses data file that consists of student test data from several districts and special schools (e.g., charter and lab schools). WVDE staff ensures (1) the file format and data elements match the final data file layout and specifications that have been approved by the state and (2) the scoring data match the processing rules/specifications.			P D	Preliminary data file, file layout, processing rules/ Specifications document
<input type="checkbox"/>	<u>After scoring MC items</u> To ensure accuracy of the answer keys and item-to-standard/objectives data in the data file	3B. The accuracy of answer keys and item-to-standard/objectives data for multiple-choice items in the preliminary data file are verified. WVDE staff scores and checks the answer keys and standard sub-score in the data file by using an independent statistical program (e.g., SPSS). The results are compared with the information on the test maps to ensure that answer keys and item-to-standard data are applied accurately by the vendor during the scoring process.			P D	Preliminary data file, test maps
<input type="checkbox"/>	<u>After scoring MC items</u> To ensure ALL multiple-choice items on each test are captured	3C. Using the preliminary data file, WVDE staff verifies the accuracy of the score range for multiple-choice items on each content area test since different tests have different numbers of multiple-choice items. The SPSS program is used for running frequency distributions of raw scores for multiple-choice items for each test. The results are compared with the information on the test maps.			P D	Preliminary data file, test maps
<input type="checkbox"/>	<u>After scoring MC items</u> To detect possible errors with multiple-choice scoring	3D. P-value or point biserials for each multiple-choice item are the indicators used in checking if any item has been scored incorrectly. Prior year operational test data for anchor items and field test data for non-anchor items are compared with preliminary current year data to assess the reasonableness of current student test results.			P D	Preliminary data file, test maps, prior year data or field test data

4. Student open-ended responses are accurately scored pursuant to specifications (scoring rubrics).

Check √	Stage/ Task Purpose	Task	Recommended Expertise*				Recommended Materials
<input type="checkbox"/>	<u>Before handscoring open-ended (OE) items</u> To ensure the training of scorers is conducted appropriately	4A. WESTEST contains open-ended items. WVDE staff (1) verifies the rules the vendor uses to hire readers are followed, (2) reviews readers' qualifications, (3) approves training procedures, materials, and rubrics the vendor uses to train handscoring readers, and (4) oversees the training and scoring process with on-site observations, conference calls, and review of status reports.	C	A			Reader qualification materials, scoring rubrics, training materials
<input type="checkbox"/>	<u>During handscoring OE items</u> To ensure the hand scoring process is conducted appropriately	4B. WVDE requires the vendor to deliver reader reliability reports on a daily basis during the hand scoring process. The procedures designed to establish the reliability of the read-behind papers are applied using random sampling (invisible read-behinds). These reports are reviewed carefully to ensure that open-ended items are scored accurately and the reliability of the readers meets the criteria. If any unusual data are observed, WVDE requires the vendor to investigate the issue and resolve the problem. Procedures executed by the vendor to ensure quality of hand scoring include: Read Behinds, Check-sets, Daily Monitoring Team, Inter-rater reliability, Daily Team Production Reports, Data Analysis, and Quality Feedback. Scorers must maintain scoring standards based upon the above procedures. Their scoring documents are audited by the vendor and must be retrained successfully or dismissed from the state's project.	A		P		Scoring rubrics, reader reliability reports
<input type="checkbox"/>	<u>After handscoring OE items</u> To ensure hand-scored data are accurate	4C. Students' scores for the open-ended items in the preliminary data file are used to verify the reasonableness and accuracy of the scored data. WVDE staff checks the score range for each open-ended item and compares the results to the scoring rubrics, hand scoring criteria/rules, and information on the test maps. Prior year data are compared with preliminary current year data to assess the reasonableness of current student test results.			P	D	Scoring rubrics, handscoring rules, test map, preliminary data file
<input type="checkbox"/>	<u>After handscoring OE items</u> To detect possible errors with hand-scored data.	4D. The percentage of students at each score point for the open-ended items is a good indicator for checking that items are scored correctly. Prior year data are compared with preliminary current year data to assess the reasonableness of current student test results.			P	D	Preliminary data file, prior year data

5. Student scored multiple-choice records and scored open-ended records are correctly matched to create a complete student record.

Check √	Stage/ Task Purpose	Task	Recommended Expertise*			Recommended Materials	
<input type="checkbox"/>	<p><u>During creating complete student records</u></p> <p>To ensure scored MC records and scored open-ended records are correctly matched</p>	<p>5A. WESTEST is composed of multiple-choice and open-ended items. WVDE requires the vendor to provide status information regarding the match of scored multiple-choice records and scored open-ended records for each student. All the open-ended items for students who returned their test booklets should be scored and matched to the students' scored multiple-choice responses. Any "questionable" matches and "unmatched" student records are examined and resolved. When necessary, WVDE staff checks with school personnel and verifies the accuracy of student records.</p>		A	P	D	Preliminary data file, tests, test maps
<input type="checkbox"/>	<p><u>After creating complete student records</u></p> <p>To ensure each tested student has a complete record</p>	<p>5B. WESTEST is composed of multiple-choice and open-ended items. WVDE staff verifies that all students in the preliminary data file (WVEIS) has both sets of scores. If one set of scores is completely missing for any student, further investigation is conducted. When necessary, the vendor must verify the accuracy of student records by reviewing the actual student test documents.</p>		A	P	D	Preliminary data file, tests, test maps

6. Scoring sheets, including item parameter tables, are generated using accurate scored student response data.

Check √	Stage/ Task Purpose	Task	Recommended Expertise*			Recommended Materials
<input type="checkbox"/>	<p><u>During equating and scaling</u></p> <p>To ensure the equating/scaling process is accurate</p>	<p>6A. WVDE requires the vendor to provide various statistical data (ICC and TCC) for review. If items are deleted for the generation of scale scores, WVDE staff reviews information regarding each deletion and also ensures that item parameter tables are revised to reflect any deleted items.</p>	A		P	Calibration data
<input type="checkbox"/>	<p><u>During equating and scaling</u></p> <p>To ensure accuracy of the calibration sample data file</p>	<p>6B. WVDE staff verifies the accuracy of the calibration sample the vendor uses to generate scoring tables. The calibration sample data file is used to (1) verify the number of item responses (score points) for each test against the expected number of item responses on the test maps, (2) review multiple-choice and open-ended item frequency distributions for reasonableness, and (3) review number of students achieving the highest and lowest possible score points for reasonableness. Unusual results are investigated and resolved prior to the production of the quality check file.</p>	A		P	Test maps, calibration data file
<input type="checkbox"/>	<p><u>During equating and scaling</u></p> <p>To ensure accuracy of the equating process</p>	<p>6C. WVDE staff and the vendor checks the accuracy of the equating procedures (CTB Itemwin Software). When possible, WVDE staff duplicates the equating process with the same software that the vendor uses to create the scoring tables.</p>			P	Calibration data, Itemwin
<input type="checkbox"/>	<p><u>After equating and scaling</u></p> <p>To ensure accuracy of item parameters</p>	<p>6D. WVDE requires the vendor to submit final item parameters generated based on current calibration sample data on anchor items and standard errors of measurement data. The data is reviewed for reasonableness and should not exceed expectations. WVDE staff compares the current year item parameters to the item parameters for those items generated in previous test administrations to ensure that there are no unexplained variances.</p>	A		P	Calibration data, scoring tables, item parameters

7. Student raw scores are accurately converted to scaled scores.

Check √	Stage/ Task Purpose	Task	Recommended Expertise*			Recommended Materials
<input type="checkbox"/>	<u>Before converting raw scores to scale scores</u> To verify accuracy of scoring tables	7A. WVDE staff verifies the accuracy of the information on the raw score to scale score tables. The current year scoring tables are compared to the scoring tables for previous test administrations. If any unusual discrepancies are observed, WVDE staff requires the vendor to double check the equating and scaling process and also conduct a meeting with the company's psychometrician(s) to investigate possible mistakes.	A		P	Raw score to scale score table, prior year tables
<input type="checkbox"/>	<u>After converting raw scores to scale score</u> To verify accuracy of the converting process	7B. WVDE staff recalculates student total raw scores for a sample of students using each student's scored multiple-choice responses and scored open-ended responses. These total raw scores are programmatically converted to the scale scores based on the scoring tables. The results are compared to the scale scores generated by the vendor to verify that the correct scoring tables are being used. Any discrepancies are indicative of possible violations of data integrity and are resolved prior to generation of test results.			P	D Preliminary data file, scoring tables
<input type="checkbox"/>	<u>After converting raw scores to scale score</u> To detect possible errors with the equating data	7C. WVDE staff conducts analyses of achievement level frequency distributions or passing/failing rates and then compares the results with prior test data to assess the reasonableness of the preliminary current test results. Any unusual differences are investigated. WVDE staff checks the score range of achievement levels (or performance standards).	A		P	D Preliminary data file, prior year data
<input type="checkbox"/>	<u>After converting raw scores to scale score</u> To ensure correct norms sets are applied	7D. WVDE staff ensures that the correct norm sets are used if using the norm-referenced tests.		A	P	D NRT tables

8. Results are accurately summarized, printed, and reported at the student, school, district, and state levels.

Check ✓	Stage/ Task Purpose	Task	Recommended Expertise*			Recommended Materials
<input type="checkbox"/>	Before producing final reports To ensure accuracy of the data file used to produce final reports	8A. WVDE staff checks ALL the data elements in the preliminary data file (WVEIS) to ensure that no data are missing, processing rules are applied correctly, and the value range for each data element is accurate.	A	P	D	Preliminary data file (WVEIS), file layout, processing rules/specs
<input type="checkbox"/>	Before producing final reports To detect possible errors with the summary reports	8B. WVDE staff compares preliminary current year data with prior year data at the school, district, and state level to assess the reasonableness of the current summary test results and passing rates. Any unusual differences are investigated.	A	P	D	Preliminary data file and sample reports, prior year reports
<input type="checkbox"/>	Before producing final reports	8C. WVDE staff verifies the accuracy of any disaggregated data before the final reports are produced.	A	P	D	Preliminary sample reports
<input type="checkbox"/>	Before producing final reports To verify the void students' scores	8D. If a student score is voided or converted to the lowest score due to test security violation (e.g., plagiarism, erasure analysis) or other circumstances, the student's score on the final file and reports are checked.	A	P	D	Preliminary data file, void student list
<input type="checkbox"/>	Before producing final reports To verify accuracy of information on the individual student reports and summary reports	8E. WVDE staff checks each piece of information printed on the sample of student reports and school/district/state summary reports against the data on the data file. An electronic comparison replaces, where possible, the manual comparison to reduce the risk of human error. PC-based statistical software (e.g., SPSS) is used to produce the same data as shown on the reports. Data from the SPSS program should match the reports produced by the vendor. Any discrepancy may indicate a possible violation of data integrity and is resolved immediately prior to generation of final test reports.	A	P	D	Preliminary data file, preliminary sample reports
<input type="checkbox"/>	Before producing final reports To verify accuracy of information on the individual student reports and summary reports	8F. WVDE staff and CTB/McGraw Hill staff conduct a 'Red Team' review to checks each sample of student reports and school/district/state summary reports. Any discrepancy is resolved immediately prior to generation of final test reports.	A	P	D	Preliminary data file, preliminary sample reports
<input type="checkbox"/>	After producing final reports To ensure the printing quality for the final reports	8G. WVDE staff conducts final reviews of variance reports between the time the reports are printed and released for shipment to ensure printing quality and accuracy of student, school, district, and state reports, including all test scores, text, and titles.				Final reports
<input type="checkbox"/>	After receiving erasure analysis and GRT files analysis will be conducted to detect possible anomalies in student scores	8H. WVDE staff compares preliminary current year data with prior year data at the school, district, and state level with the erasure analysis to assess the reasonableness of the current summary test results and passing rates. Any unusual differences are investigated.	A	P	D	GRT data file and erasure analysis, prior year reports

* Key for Professional Expertise Recommendations: C = Content A = Assessment P = Psychometric/Research D = Data Management