Ohio's Machine-Scoring Process

Step 1: Field Testing Field test items are embedded in student assessments. These items do not count towards a student's score on the assessment, however they provide the information necessary to identify the range of responses that could be expected when this item is included in future operational assessments. Additional information on the development and use of test items in Ohio can be found here.

Step 2: Range Finding Ohio educators review the student responses provided to the field-tested items. Educators select student responses that represent a full range of possible answers at each score level. These samples will be used to create guidance to train scorers.

Step 3: Hand Scoring of Field Test Items Ohio's scoring contractor, Data Recognition Corporation, uses student responses selected during the range-finding process to train hand scorers. Scorers are provided a rubric and multiple sample responses that represent each possible score level. They then score the entire pool of field test responses.

2,500 randomly selected responses, representative of Ohio's student population, are then hand scored a second time to validate the human scoring process. These validated responses are then able to be used to program the scoring engine. During this process, every discrepancy between scores is resolved by a third human scorer. The quality of the human-assigned scores is critical to the identification of a valid model and final performance of the scoring engine.

Step 4: Programing the Scoring Engine The testing contractor, AIR, then begins the process of training the scoring engine based on 2,000 randomly selected responses of the 2,500 that have been scored and validated. Additional information on the development and use of test items in Ohio can be found here.

Step 5: Quality Control Adjustments The machine scoring engine is adjusted through multiple quality checks to confirm that the sample pages are being scored appropriately.

Step 6: Verification of Machine Scoring Accuracy From the sample of 2,500 responses, we reserve approximately 500 responses for an independent validation. AIR verifies that the machine-scoring engine scored the student responses the same as the human scorers. If discrepancies are identified, the scoring engine is adjusted (see Step 5).

Step 7: Scoring the Statewide Operational Test Once AIR has a verified scoring engine and is programmed to score appropriately, the machine is ready to score the statewide operational test. AIR and the Data Recognition Corporation have built-in verifications to make sure the machine is performing accurately.
- Verification one: The first 500 responses submitted during the testing window will be machine-scored and then human-scored to check scoring validity and reliability.
- Verification two: This occurs when 25 percent of the student responses throughout the test window are also machine scored and then hand scored for verification.

Step 8: Ohio Department of Education Review All scoring procedures and results are reviewed with AIR and with Ohio's Technical Advisory Committee, made up of national assessment experts.

Step 9: District Review Districts receive their results once uploaded into the Online Reporting System. Districts have the option to request a rescore of student responses.

Updated March 2018