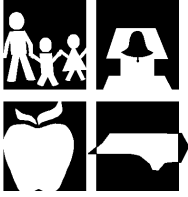


This publication and the information contained within must not be used for personal or financial gain. North Carolina LEA school officials and teachers, parents, and students may download and duplicate this publication for instructional and educational purposes only. Others may not duplicate or quote from this publication without prior written permission from the NCDPI Division of Accountability Services/Testing Section.



North Carolina Writing Assessments: A Procedural Outline of Prompt Development, Field Testing, Prompt Selection, and Scoring Process*

Public Schools of North Carolina
The North Carolina Department of Public Instruction
Division of Accountability Services/Testing Section

Prompt Development and Selection

At each grade level, prompts for field testing are developed by members of an Advisory Committee consisting of North Carolina educators, administrators, and North Carolina Department of Public Instruction (NCDPI) curriculum and assessment staff. This committee then reviews each prompt for curricular appropriateness, potential bias, and language/age appropriateness. Prompts that meet these criteria are then selected for field testing with a random representative sample of North Carolina students.

During field testing, North Carolina teachers are asked to complete a form rating each prompt to which their students responded. This form includes four rating categories: Curricular Appropriateness, Language Appropriateness and Clarity, Format Quality, and Overall Appropriateness and Quality, as well as space for additional comments. These data are collected, entered into a database, and later reviewed.

The field-test responses are then scored by a contractor who has previously scored a North Carolina statewide administration at the same grade level. The contractor's scoring director reports on the performance of each prompt, focusing on such points as any miscues created by the wording of the prompt, appropriateness of the topic for the level assessed, and an overall impression regarding the suitability of the prompt for statewide administration. Statistical data are also collected on overall score point distributions, along with the performance of various demographic groups (e.g., ethnicity and gender) on each prompt.

All of the above field-test information and data are carefully reviewed by NCDPI curriculum and assessment specialists, and again by the advisory committees, before a prompt is given final approval for statewide administration.

Range-Finding and Selection of Training Materials

Immediately following statewide writing assessments, the NCDPI advisory committee for the appropriate grade level meets to discuss and score a representative selection of student field-test responses to a prompt. These committees are comprised of North Carolina educators and administrators, as well as curriculum and assessment specialists from the NCDPI and North Carolina State University's Technical Outreach for Public Schools (TOPS). The contractor's scoring directors also attend these meetings. Each member of a committee comes to the meetings having first scored the papers independently.

During this initial scoring process, known as range finding, student responses are discussed and assigned "true" scores. These scores are derived from careful consideration of the following: the four scoring criteria that are applied to all NC writing assessments (main idea, supporting detail, organization, and coherence); the specific scoring rubric that applies to each grade and mode of writing; and the student papers from previous years' scoring guides which serve as "anchors" to ensure the consistent application of standards from administration to administration.** From these range-finding papers, responses that represent the spectrum of all score points are selected for the scoring guide, training sets, and qualifying sets used in training the team leaders and readers who will score the student responses.

* The procedures described in this document may vary slightly depending upon the grade level of the assessment, and are similar to those used for the North Carolina Open-Ended assessments at grades 4 & 8.

** The scoring criteria and rubrics are discussed in the Scoring Guides released after each administration. These guides may be accessed at <http://www.dpi.state.nc.us/accountability/testing/index.html>.

Overview of Contractor Scoring

As is the case with most state education agencies that conduct statewide testing programs, the NCDPI contracts for the professional scoring of student papers. There are several advantages to using a contractor specializing in performance assessment scoring, including: proven reliability and validity of statewide assessment scores; daily and project-long data reporting and monitoring capabilities; experienced logistical, security and systems analyst staff; consistent supervisory and training staff from one administration to another; and the availability of a pool of qualified, experienced readers for the three to six week duration of a scoring project.

The contractor is responsible for receiving the student responses, separating the responses from the student and school information portion of the test booklets, recording a unique identification number on each paper, and placing papers in packets of approximately 20 for scoring. Two scannable monitor sheets for recording scores are generated for each packet, as each student response is read independently at least twice. These monitor sheets allow for daily statistical monitoring of individual readers, groups of readers, and the project as a whole.

The contractor provides for strict security throughout the scoring process. Readers must sign a confidentiality agreement, wear an identification badge, and leave all scoring and training materials in the scoring rooms. Security personnel monitor compliance with all security procedures.

The contractor is also responsible for providing qualified staff, including scoring directors and readers, who meet the requirements of the NCDPI. Readers are organized into teams that are coordinated by team leaders. The contractor ensures the consistent and effective training and monitoring of team leaders and readers throughout the project. The contractor must meet all of the NCDPI standards for a project, including the scoring accuracy criteria established by the North Carolina State Board of Education.

The NCDPI monitors all contractor training and scoring at every step of the process. The papers used in scoring guides as well as training materials have been pre-scored and approved by a NCDPI Advisory Committee. NCDPI staff are present on-site during the training of team leaders and readers. During actual scoring, NCDPI staff are in constant contact with the contractor, discussing papers and monitoring statistical data daily.

Contractor Staff Qualifications

Scoring Directors

- Scoring Directors are selected by the contractor and approved by the NCDPI staff. These directors are experienced in scoring a variety of statewide writing/performance assessments and have proven their ability to train team leaders and readers effectively, monitor all facets of scoring, and ensure the consistent and accurate application of state scoring criteria. For example, the content scoring director for the 1999–2000 English II administrations holds a bachelor’s degree in English Education with a minor in reading. This director also has over four years of professional scoring experience with several statewide performance assessments and has served as the content scoring director for the focused holistic portion of the North Carolina English II End-of-Course Test for the past two years.

Team Leaders and Readers

- Requirements for readers and team leaders include a minimum of a bachelor’s degree in an appropriate academic discipline, e.g., English, Education, Language Arts, etc.; teaching experience is preferred. In addition, all potential team leaders and readers are required to successfully complete a personal interview, a writing assignment, a scoring screening sample, and a grammar test. Candidates are also required to demonstrate editing ability.
- Team leaders are selected from the pool of readers on the basis of demonstrated expertise regarding all facets of the scoring process, including strong organizational abilities and skill in training strategies.

Team Leader and Reader Training

All team leaders and readers receive extensive training and are required to meet the NCDPI standards for demonstrating scoring accuracy prior to scoring actual student papers. The first objective of training is to remove, at least temporarily, any biases that readers might hold concerning the related importance and appropriateness of certain characteristics of written composition. The training materials are designed to give clear definition to each characteristic that the reader will be asked to evaluate.

NCDPI staff are on-site to monitor the entire training and qualifying process for both team leaders and readers.

Team Leader Training

The scoring director trains team leaders for several days prior to the beginning of reader training. The guidelines and standards for training team leaders are essentially the same as for readers outlined below. Team leaders must completely understand and annotate all training materials so they are prepared to assist in explaining, interpreting, and analyzing responses during reader training.

Reader Training

Reader training begins with an overall introduction to the assessment both by the scoring director and by NCDPI staff. The scoring director then introduces and discusses the NCDPI focused holistic scoring criteria in conjunction with the “anchor papers” in the scoring guide. This allows for readers to understand the scoring criteria in the context of specific papers. The scoring guide includes at least three papers representing each score point, and these papers have been annotated, using the language of the scoring criteria. (As with all training materials, the “true” scores for these papers are assigned by the NCDPI and its Advisory Committee. See the “Range-Finding and Training Materials” section above.)

After presentation and discussion of the scoring guide and anchor papers, each reader is given a training set of 10 papers to score independently. Once the readers have completed scoring these 10 papers, the scoring director begins discussion of the set by reporting the “true” scores assigned to the papers. This affords readers the opportunity to discover how similar or different his or her approach is to that of the NCDPI and its Advisory Committee. The scoring director then leads the discussion of each paper, focusing on the rationale behind the assigned “true” scores. Group discussion is encouraged, as this allows readers to raise questions about the application of the scoring criteria to specific papers. Readers then score independently another four training sets consisting of a total of 55 papers. After each set, the scoring director or team leaders discuss the “true” scores of each paper. Throughout this process, reader scores are recorded and reviewed in order to identify both the scoring trends of individual readers and any papers that may be problematic for a group of readers.

Upon completion of the training sets, readers are required to score 3 qualifying sets of 20 papers each. A reader must achieve a certain level of perfect agreement with the “true” scores on at least one qualifying set before being allowed to score actual papers. For grades 4 and 7, that percentage is 75%. If a reader candidate does not meet this standard, he or she is dismissed from the project.

Ongoing Reader Training and Monitoring

After readers are qualified and the actual scoring process begins, several methods are employed to ensure consistent and continued scoring accuracy by all readers, including the following:

- At least seven sets of 20 pre-scored actual student papers are assembled in “validity packets” that are indistinguishable from the rest of the unscored packets. The NCDPI and the scoring director have agreed upon the “true” scores of all validity papers. These validity packets are one method used to detect whether readers are continuing to accurately apply the scoring criteria. These packets are circulated daily among the readers and scored by them as any other packet is. When an individual reader has scored a validity packet, that reader’s scores are immediately scanned and compared to the “true” scores. Reports of these “validity” scores are generated and closely checked by the scoring director. The reports are compiled and forwarded daily to the NCDPI for review.

- Team leaders re-read a random sample of approximately 4 percent of papers scored by readers on his or her team to ensure team members are correctly applying the scoring criteria. If a team leader identifies a reader who is not consistently applying the criteria, the team leader will discuss these issues with regard to specific papers to correct the problem. (NOTE: A reader may be dismissed from the project at any time if his or her scoring does not meet the NCDPI standards.) Team leaders also “cross score” with other team leaders, i.e., two team leaders will score the same packet independently. Comparison of these scores further ensures consistency across readers.
- Readers record their scores on scannable monitor sheets that allow for reliability reports to be generated at least twice daily during scoring. These inter-reader agreement reports (for both individual readers and groups of readers) include the number of papers scored, the percentage of papers assigned each of the score points, and the exact agreement data. The percentage and number of adjacent (one score point difference) and resolution (more than one score point difference) scores are also included. These data are used to identify any possible “drifts” in applying the scoring criteria and allows for immediate correction on the part of team leaders or the scoring director. These reports are sent daily to NCDPI for review.
- Readers and team leaders are trained to “flag” papers that are unusual or that may call for additional consideration. Unusual responses or approaches are brought to the scoring director and reviewed with NCDPI assessment staff.
- Scoring directors will often begin the morning by refocusing the group on the scoring criteria and specific scoring guide or training set papers that clearly illustrate specific score points.

Inter-Reader Reliability Agreement and Overall Reliability of Scores

Each group of readers is trained on only one grade level of the North Carolina Writing Assessments. Every student response at grades 4, 7 and 10 is scored independently by two readers. Each reader’s scores are recorded on a scannable monitor sheet and the data are analyzed to produce inter-reader reliability reports throughout each scoring project. The minimum standard for inter-reader reliability, established by the North Carolina State Board of Education, is 70 percent. Traditionally, the scoring contractor has exceeded that standard. For example, during the 1998–1999 grade 7 writing assessment, the overall perfect agreement rate was 79.7 percent. The adjacent agreement rate (e.g., a paper was assigned a score of a “3” and a “4”) was 20.2 percent of the total 90,921 papers scored.

The relatively low percentage of adjacent scores reflects the fact that most papers clearly fall within a specific score point. However, there are instances when two readers, both of whom have consistently demonstrated scoring accuracy and understanding of the scoring criteria, assign adjacent scores to a particular paper. Based on the scoring criteria and sample student responses that illustrate the specific score points, one reader may determine that a paper more accurately compares to a high “3,” while the other reader may determine that the paper more accurately compares to low “4.” In such a case, a designation of “3.5” most appropriately defines the score of that paper.

Non-adjacent scores (e.g., papers that are scored by two readers with score point of a “2” and a “4”) are rare, approximately 0.1 percent. These papers are read a third time by the scoring director who then assigns the final score.

(Note: A process is in place for appealing a student’s score to the NCDPI after careful review of the student’s response at the LEA level.)