

NEW YORK STATE TESTING PROGRAM

SCORING LEADER HANDBOOK

Grade 8 - Listening, Writing / Writing Mechanics, and Reading
March 2002

The scoring leader position is a very important part of the New York State Testing Program. The success of this scoring project depends on the scoring leaders' understanding of the scoring criteria and on their ability to explain these criteria to the scorers they train. Each scoring leader is responsible for creating a comfortable, professional environment while also setting a productive pace for the training and scoring session(s). The scoring leader will need to establish a dialogue with the scorers while at the same time maintaining the pace of training. In addition, each scoring leader should set a positive tone during training by conveying confidence in, and support of, the overall merits of the test itself and the process by which scoring decisions were made. Each scoring leader will need to be prepared to answer questions with patience and diplomacy, keeping in mind that the overall goal is to train the scorers quickly to score accurately and confidently.

Preparation for Training

In order to maintain scoring consistency from site to site, it is crucial that all scoring leaders explain the training materials using the same language and emphasis. You will need:

1. a Scoring Guide that contains the scoring rubric, exemplars (student responses that illustrate each score), and printed annotations (information that explains the scoring decision for each exemplar); and
2. a Scoring Leader Practice set that consists of approximately 10 student responses, in each content area—Listening, Writing, and Reading. The Writing Mechanics Practice Set consists of approximately 5 student responses.

Note: For each content area, the Scoring Leader Practice Set also contains margin notes pointing out important aspects of the student responses in relation to the scores they received. (The scorers' training materials are the same as the scoring leaders' materials except that the Scorer Practice Set responses do not contain the annotations, margin notes or the scores because they are to be used for scoring practice.)

3. a videotape training session for either Session 1, Part 2: Listening; Session 2, Part 1: Reading, Session 2, Part 2: Writing. The tape **must** be used for training scoring leaders. The tapes may also be used by scoring leaders to train actual scorers, or the scoring leaders may choose to train scorers based on the training they have received through the use of the tapes. Permission is granted to duplicate the tapes if additional copies are desired.

Using the Videotapes to Train Scorers

In each videotape, the presenter discusses the contents of the Scoring Guide and the Practice Set for that content area. The Scoring Guide will be presented first to demonstrate how the scoring rubric should be applied to student responses. The Training Leaders should stop or pause the video  before the videotaped presenter begins discussion of the Practice Set. This provides an opportunity for those being trained to read their Practice Sets and practice making scoring decisions. The scorers should practice on only one of two student responses at a time, stopping

and reviewing the correct score(s) before moving on to the next. The scoring leader may read and discuss the annotations and marginalia in their copies of the Practice Sets, or may resume the videotape at appropriate intervals. Several short practice segments followed by review maximizes the opportunity to learn by doing and assists in building scorer skill and confidence.

While viewing the scorer training tapes you may add additional notes as needed. Any questions that arise during your training may very well come up again during scorer training; therefore, this training session is the time to prepare to answer such questions by making consistent notes on both your Scoring Guide and your Practice Set. Additional scoring support will be provided by a Question and Answer document, which will be placed on the Department's web site on March 13th. The web site address is: <http://www.emsc.nysed.gov/ciai/assess.html>

Once at the web site select Grade 8 English Language Arts 2002 Question and Answers. There will also be a toll free number that Scoring Site Coordinators or Scoring Leaders may call with questions pertaining to the training or scoring. This telephone service is available on weekdays from 8:00 a.m. – 5:00 p.m. from March 12, 2002 – March 27, 2002. The number is (877) 516-2403 and can also be found, along with the schedule and a fax number, at the end of this Handbook.

Your mastery of the scoring terminology and complete knowledge of the training materials will prepare you to conduct scorer training successfully. You should know the scoring rationale for all the exemplars and be prepared to answer any questions about the scoring decisions using the appropriate terminology from the rubric.

Rehearsing your delivery prior to the training day will be helpful. Practice reading the rubric, exemplars, and annotations out loud; then practice using them and the handwritten notes to help  scorers understand the exemplars. Remember that you want your presentation to be fresh and interesting rather than just a mechanical reading of notes. A thorough understanding of and familiarity with the training materials will prepare you to “think on your feet,” and successfully answer any questions that may arise during scorer training.

Training

If you do central scoring:

Each scoring site will have a site coordinator and scoring site assistants (clerical aides). You should meet with the site coordinator prior to the day of training and scoring in order to:

1. learn your site's paper flow plan for scoring the student responses,
2. make sure that your scoring/training area is ready,
3. ensure that all necessary materials are on hand and that they are counted, and
4. discuss any other logistical issues.

Following is a suggested schedule for the NYSTP training/scoring day:

| | |
|--------------------|----------|
| 8:40 - 11:20 a.m. | Training |
| 11:20 - 12:20 p.m. | Lunch |
| 12:20 – 3:00 p.m.. | Scoring |

Suggested break times are 10:00-10:15 a.m. and 1:45-2:00 p.m. These times may change if needed; consult with the site coordinator. It may be necessary to stagger the break times by room, depending on the number of scorers at your site and the availability of restroom facilities.

Ideally, the site coordinator will ask the scorers to arrive at 8:15 a.m. to take care of sign-in, nametags, and other logistical issues. The site coordinator will be responsible for greeting scorers and directing them to your room; you should be in your area no later than 8:15 a.m. to meet your scorers and make sure you are ready to begin training promptly at 8:40 a.m.

Prior to 8:40 a.m., you should distribute the materials to the tables or desks where the scorers will be sitting. (The site coordinator will have your room and tables ready for you.)

Each scorer should have: Scoring Guide
Practice Set
Sharpened No. 2 scoring pencils, pens (for use during training),
Post-it notes (yellow flags), erasers, etc. (provided by site coordinators)

At 8:40 a.m., all your scorers should be present and seated, and training can begin.

If you score locally:

Each scoring site should have a school administrator who will be in charge of scoring operations, including supervising the scoring and scoring operations, coordinating test booklet processing, identifying support needs, sending answer documents to the scanning center, and enforcing security. You should meet with the administrator prior to the day of training and scoring in order to:

1. learn your site's paper flow plan for scoring the student responses,
2. make sure that your scoring/training area is ready,
3. ensure that all necessary materials are there and that they are counted, and
4. discuss any other logistical issues.

If you like, you may follow the same training schedule suggested for central scoring. You should be in your area no later than 8:15 a.m. to welcome your scorers and make sure you are ready to begin training promptly at 8:40 a.m. Prior to 8:40 a.m. you should distribute the materials to the tables or desks where the scorers will be sitting.

Each scorer should have: Scoring Guide
Practice Set
Sharpened No. 2 pencils, pens (for use during training),
Post-it notes (yellow flags), erasers, etc. (provided by the school administrator)

At 8:40 a.m. all your scorers should be present and seated, and training can begin.

Suggested Training Agenda

The suggested Training Agenda and detailed procedures outlined on pages 3-11 provide useful information for all scoring leaders, but are essential if the tapes are not used to train scorers. Those who choose to use the tapes to train scorers will find that most of this information is contained on the tapes.

1. Introduce yourself.
2. Introduce others, if applicable (State Education Department representatives, site coordinator, site assistants, table facilitators).
3. Review housekeeping details (the day's schedule, break times, restroom and drink machine locations, lunch location, smoking areas, etc.).

Once the above issues have been addressed, you will be ready to move into the actual training.

4. Briefly define holistic scoring and the scorer's responsibilities.
Tell your scorers that you will be training them to use a process called holistic scoring. This type of scoring involves evaluating a student's work for its total, overall, or whole effect based on the rubric and accompanying exemplar responses. In the New York State Testing Program, this process involves cluster scoring, where multiple tasks are evaluated and scored as a whole presentation; i.e., multiple tasks are given one overall score. Cluster scoring provides a more authentic measure of a student's skills than does item specific scoring; however, cluster scoring is demanding in that scorers must keep in mind the responses to all of the items in a cluster while applying a holistic rubric.

Emphasize that holistic scoring is similar to learning a new language or a new way of thinking, and that it is crucial that all scorers put aside their own beliefs, ideas, and theories about how to evaluate students' work. For any large scale scoring project to be successful and have meaningful results, all scorers must score using the established criteria. Therefore, you will be training your scorers to understand and internalize the NYSTP criteria. You will do this by explaining the rubric, along with student exemplars for each score. Assure your scorers that the more exemplars they see, the clearer the criteria will become.

Tell your scorers that this training session is not the time to critique the test questions, the rubric, or the scoring decisions. **The purpose of this training is to learn to apply the scoring criteria, not to give opinions of how to alter the test or the criteria.** Inform them that committees of teachers assembled by the New York State Education Department made all scoring decisions. Let them know that any insights or opinions they have about the criteria may be sent to the State Education Department. Explain that there is a lot of training material to cover and not much time to do so.

Scoring Guide (Suggested Training Time: 1 hour and 20 minutes)

At this point, you should ask the scorers to access their Scoring Guide. Explain that the Guide contains exemplars that will be used as references or anchors when the scorers begin to score live test booklets. Ask them not to read ahead, but rather to stay on the page you are discussing.

As an overview, tell your scorers that the Guide (except for Writing/Writing Mechanics) consists of the following: the Reading or Listening passages, two rubrics (a generic rubric for Grade 8 English Language Arts and a specific rubric tailored to the actual test items to be scored), Rubric Key Points (to be explained later), exemplar responses, the ELA Scoring Considerations, the ELA Condition Codes, and both the generic and specific rubrics in chart form.

Whenever you are introducing material in the Scoring Guide, you should inform the scorers which page you are on so that they can read along silently while you are reading aloud. This will keep the group (who may read at different speeds) together. This process also helps the scorers to internalize the criteria, because they are simultaneously reading and hearing the information. You should read all the material in the Scoring Guide (the rubric and exemplars) aloud. You should also encourage the scorers to take notes on their materials during training. Emphasize that these materials are theirs to use during the scoring session but the scoring materials must be stored in a secure location in their school until the date indicated in the School Administrator’s Manual, and also noted at the end of this Handbook.

Begin the training by introducing each of the specific assignment tasks the students were given, as well as any accompanying resource material, if applicable. For example, in Listening, students heard and in Reading they read passages before they answered the test questions. All of this material is in the Guide. So first read the passages and then the questions aloud (using the first exemplar response) aloud while the scorers read silently.

Reading and Listening only: Next, turn to the generic rubric, called the Grade 8 English Language Arts Rubric, and briefly review it. The scoring scale goes from 0-6. Tell your scorers that a score of 0 is assigned to the lowest performance level holistically and that the highest score point is reserved for the best performance level. Point out that the same rubric—in chart form—is at the back of the Guide for quick reference.

Following the generic rubric is your content area’s specific rubric, which describes all score points in relation to the particular test items. Do not review the entire range of score point descriptions at this time. Rather, simply note its possible use later as quick reference and move on to the Rubric Key Points.

The Rubric Key Points are designed to show some of the kinds of responses that would be acceptable for each item in the cluster. Point out that these examples are not exhaustive. Your scorers must realize that more than a quarter of a million eighth-grade minds have responded to these test items, and some of those minds might have come up with perfectly sound, relevant text-based responses that have not been anticipated by the adults who developed the Key Points.

Make clear to the scorers that they are not required to use this instrument in their scoring. If they choose to use it, they must always keep in mind that “other relevant text-based responses” are acceptable.

Writing has only one rubric.

The exemplars in your Scoring Guide are formatted to progress from the lowest score to the highest. The specific rubric description of score point 1 precedes the exemplars of score point 1, followed by the score point 2 description and accompanying exemplars, and so forth. Before beginning to discuss the 1's, however, turn ahead in the Guide to read an exemplar that earned the highest score—a 3 (Writing/Writing Mechanics) or 6 (Listening, Reading)—so the scorers will understand where you are going as you move up the score point scale. Then turn back to the 1's in the Guide.

Read and explain the rubric for score point 1 and then read and discuss each annotated exemplar. Answer any questions and then move on to score point 2, and so on. Occasionally, a score point 0 exemplar is included in the Guide, in which case, you should read and discuss the 0 before you discuss score point 1. If no 0 exemplar is included, make sure to still read and discuss the rubric description of a 0.

Throughout, maintain an atmosphere that promotes clarification rather than debate. Do not let a discussion become contentious and therefore counterproductive. If a scorer absolutely refuses to see a certain exemplar as the score it is, advise him/her to forget the exemplar, but learn the lesson it demonstrates about the characteristics of the score point in question. There are multiple samples of each score point in the Guide and training sets so that the scorers will have ample exemplars to use to help them make good scoring decisions. One or two “controversial” exemplars should not derail the training process or prevent you from training the scorers to score accurately.

Answer questions patiently and thoroughly, but feel free to say, “It’s time to move on,” if you feel that the discussion is starting to deteriorate. Part of the scoring leader’s job is to maintain control of the group.

It is helpful to demonstrate the use of yellow flags (Post-it Notes) to your scorers. Use these flags to index the Scoring Guide on each page where a score point is introduced. (Attach a flag with a 1 to the upper right side of the specific score point 1 description; then attach a flag with a 2 on the right side of the specific score point 2 description but slightly lower than the 1 flag, and so on. Place each flag slightly lower so that all score point numbers can be easily seen.) This way, during the scoring/discussion of the training sets and the scoring of students' test responses, the scorers can easily access a score point exemplar for reference and comparison.

Explain that accurate scoring comes from using the Scoring Guide effectively—the rubric description for a particular score point should always be referenced in conjunction with the exemplars for that score point. The exemplar responses act to elaborate upon the rubric and help the scorers to interpret them correctly. The student exemplars can be used effectively for reference and comparison.

The next-to-last page in the Scoring Guide addresses condition codes and scoring considerations. Condition codes are letter codes assigned to student responses that are not scorable. Scoring considerations are guidelines that are specific to the tasks you are scoring, and address issues such as what to do when one of the tasks in the cluster is blank. There are no examples of the condition codes or the scoring conditions in the training materials. Tell your scorers that you will explain these applications after completing discussion of the Guide and Practice Set.

Practice Set (Suggested time: 40-45 minutes to score and discuss)

Once you have completed the discussion of the Scoring Guide, move directly to the Practice Set. Have scorers access the set. Explain that this set is an opportunity for them to practice scoring; they should use the criteria they have internalized from the discussion of the Scoring Guide to score the student responses on their own.

Tell them how many student responses are in the set, and that the responses are arranged in random order. (Unlike the Guide, the samples will not begin with the lowest score and move up the score point scale.) You will move through the set one response at a time. Remind them that 3-4 items make up a cluster response. The scorers are to consider the cluster as a unit and apply one score to that student's cluster response.

Tell the scorers to read the cluster of the first student response silently to themselves and write down a score on the response. Encourage them to base their score on their overall holistic impression; if their impression is “either a 2 or a 3,” tell them to reference the exemplars in the Scoring Guide to see if the Practice Set response is more like a 2 or a 3. Give them a couple of minutes to read and score the first sample, then tell them the correct score, explain the rationale for the score, answer any questions, and move on to the second Practice Set sample. Move through the entire Practice Set in this manner.

Like your Guide, your Scoring Leader Practice Set is annotated so that you are prepared to explain the scoring decisions. The Scorers’ Sets, however, will be completely unannotated, so remind them to take notes as you explain the scoring decisions. Be prepared to explain a score from both directions. For example, a sample with a correct score of 2 may have received both 1’s and 3’s from the scorers, so you should be prepared to explain why it is not a 1 and why it is not a 3.

When applicable, your notes also contain specific Scoring Guide exemplars with which to compare the Practice Set responses; the best way to justify a scoring decision is to show how the sample compares with the exemplars in the Guide or with previous Practice Set responses. You, as scoring leader, should be supportive and positive during this training process and should keep bringing the scorers back to the Scoring Guide and the anchor exemplars. Tell the scorers not to worry or agonize if they incorrectly scored several of the samples. This is a Practice Set that will introduce them to a variety of responses, some of which are different in approach from the Guide's exemplars, so they will be new to the scorers. Much can be learned from incorrectly scoring responses because one tends to try harder to understand the scoring rationale of the responses that one mis-scores. Remind the scorers that the goal is to understand why each sample

received the score it did, and that it is more productive to focus on why a paper is a 2 than to argue why it should be a different score from the one assigned.

Despite your thorough preparation, it is always possible that a scorer might ask a question for which you do not know the answer. Please feel comfortable saying, “I don’t know, but I will try to find out,” and move the training forward. You can later try to find out the information from another scoring leader at your site, or by calling the Help Line set up by MI (Measurement Incorporated). The toll free number and its hours of operations are provided at the end of this handbook.

Another type of question you should be prepared to address concerns theoretical student responses. Scorers may say, “What if the student had done this?” or “If this particular task were better, would the overall score of the cluster change?” It is recommended that you tell your scorers that you would prefer to talk only about actual student responses rather than theoretical ones, because talking about responses that don’t really exist can cause unnecessary confusion. It is safer to stick to the written responses that all scorers can be looking at while discussing a scoring decision. That way everyone will be seeing exactly the same thing.

The challenge for the scoring leader during this part of the training process is twofold: to remain diplomatic and patient if any scorers become frustrated, and at the same time, to keep the training process moving forward. You should listen to the scorers’ questions and concerns and address them as thoroughly as possible while also keeping an eye on the clock with your schedule in mind.

Grade 8 Writing Mechanics Training

The Grade 8 Writing content area contains an additional category called Writing Mechanics, which has its own Guide and a Practice Set. After scoring a student's written response to the Writing passage using the Writing rubric, the scorer will then score that student's remaining two written responses (Listening and Reading) for Writing Mechanics using the Writing Mechanics rubric. The Writing Mechanics Guide usually contains one student cluster exemplar response for each score point, from 0 to 3, although occasionally there may be no 0 exemplar. The overall response is scored for the conventions of written English (grammar, syntax, capitalization, punctuation, and paragraphing).

The Writing Mechanics Guide will contain the passages to which the students are responding. Do not get bogged down in content-related questions, because the focus at this time is mechanics, which is not content bound. The procedure for presenting the Writing Mechanics Guide and Practice Set is the same as for the Writing training materials.

Paper Flow/Scoring Procedures (Suggested Time: 25 min.)

Once training is complete, you will need to show the scorers how to score “live” test booklets. Each scorer will have been assigned a scorer number—your site coordinator should provide you with a scorer roster with these numbers on it. Make sure each scorer knows his/her number and has a sharpened pencil.

Give each scorer an unscored student test booklet and take the group through the steps of scoring a booklet. You will need a blank booklet and monitor for demonstration.

1. **Scorer numbers.** The scorer should write his/her number on the upper right hand corner of the cover of the test booklet.

Note: The scorer number will be three digits. This three digit is the individual's "scorer number", which is bubbled into a field on the answer document. Numbers may be assigned by group, for example:

| | |
|-------------------|------------|
| Listening scorers | 100 -- 199 |
| Writing scorers | 200 -- 299 |
| Reading scorers | 300 -- 399 |

A numbering system like this will enable everyone to look at the upper right hand corner of the test booklet cover to tell which groups have not scored the booklet.

2. **Answer Documents.** A variety of answer documents are being used by the different scan centers. Before training your scorers, you should find out from your site coordinator what the answer document for your site looks like so you can instruct your scorers on how to fill it out properly. Ask the scorers to remove the answer document from the test booklet and show them where they write (and bubble) their scorer number. Explain the care and handling of answer documents (do not fold, or wrinkle or make any stray marks on the sheets). Show them the area where they write their holistic score and where they bubble their score. Emphasize that these answer documents will be machine scored, so care must be taken to avoid stray marks. Also, show them how to compare the booklet number with the answer document number; this is how you make sure you are using the right answer document for a particular booklet.

Emphasize that scorers must both write the score and bubble it in!

3. **Proper bubbling.** Score items only once, make dark bubbles (don't forget the center of the bubble) and if it becomes necessary to erase, erase neatly and thoroughly. **Scorers should never write on the test booklets (except for the scorer number on the cover). Scores go only on the answer document!**
4. **The paper flow system.** The Scoring Site Operations Manual (January 1999) details the paper flow system. See your site coordinator for this information. Ideally, no scorer will have "downtime" when he/she is waiting for booklets to score.
5. **Condition codes that should be flagged.** At or near the end of the Scoring Guide are the definitions of the codes; ask the scorers to turn to this page. Scorers may not assign a condition code to a student response; only the scoring leader may assign the codes. Scorers need only to recognize the types of responses that receive a code, so that they can flag them for you.

The only exception to this is Code A = Blank. Scorers may apply Code A if the entire student response is blank. All other codes should be flagged (use the yellow post-it notes) and the flagged booklets brought to the scoring leader by the site assistant or table facilitator.

6. Other types of responses that should be flagged.

a. Sensitive papers. During the field testing conducted for the New York State Testing Program, teachers reading student essays discovered that some students wrote about sensitive issues. If a teacher reads a student response that reveals a sensitive issue, he or she should share this essay with the table facilitator and the scoring site coordinator. A sensitive response would include:

- evidence of parental or teacher abuse,
- suicidal tendencies, and/or
- other psychological problem.

The scorer should score the response according to the ordinary rules and then flag the response by writing “sensitive paper” on a post-it note and signaling the table facilitator or the site assistant. Scoring leaders need to notify the school principal of any sensitive responses. If tests are being scored regionally, scoring leaders should alert the site coordinator who will contact the student’s principal.

b. Scoring Decisions/Odd Responses. These are responses that the scorer is unsure about; e.g., none of the exemplars in the Scoring Guide help the scorer to make a scoring decision about these particular responses. The scoring leader should make a decision, bubble in the score, and return the booklet to its appropriate box.

c. Booklet Problems. (answer document is missing, answer document and booklet don’t match, etc.) You should have a Problem Box for these types of problems so that the site coordinator can handle/solve them.

Note: Instruct the scorers in the proper use of yellow flags for the above issues. The responses in question should be flagged on top of the test booklet—the flag should be easily visible and the type of problem/situation should be written on the flag (sensitive paper, insufficient, illegible, etc.). The booklet can then be put into the Problem Box. The site assistant or table facilitator will bring all flagged booklets to you. You should deal with the flagged booklets that are your responsibility (assign a condition code, make a scoring decision, etc.) as quickly as possible so that the packets can return to circulation in order to make sure that they are completely scored by the end of the day.

Don’t allow these booklets to pile up!

No flagged booklets should be transferred to another scoring room until the issue flagged is addressed!

Scoring

Once you have covered the Scoring Guide, Practice Sets, and paper flow/booklet logistics, the scorers may begin scoring!

Encourage them to score accurately and productively. You may want to give them a goal or expectation. (For example, each scorer should score at least 50 booklets during that day.) Consult the site coordinator for the specific productivity goal for the scorers in your content area.

While you do not want the scorers to feel that speed is more important than accuracy, you also want to make sure that all booklets are scored by the end of the day. Remind the scorers that the holistic scoring process is, by design, a rapid scoring process.

If you have completed the training before the lunch break, then have the scorers score until you release them for lunch at the pre-established lunch time. Encourage the scorers to return promptly at the designated time and begin “live” scoring immediately.

During the afternoon, the scoring room should be kept as quiet as possible to facilitate accurate, productive scoring. Emphasize that the scorers should discuss scoring only with the table facilitators or the scoring leader, in order to avoid slowing down production and creating a disturbance for others. Meanwhile, you should score flagged booklets, answer questions, troubleshoot problems, and time permitting, score booklets yourself.

Paper Flow from Room to Room

The site coordinator will train the site assistants in the logistics of transferring booklets from Reading to Writing to Listening. The site assistants will also be responsible for checking all answer documents to make sure that they are complete and accurate.

Measurement Incorporated

Grade 8 ELA Help Line

(877) 516-2403

**Dates: Tuesday, March 12
Through
Friday, March 27**

Hours: 8:00 A.M. – 5:00 P.M. EST

Fax #: (919) 425-7733

**Scoring materials must be stored in a secure
location in schools until March 28, 2002**

