

APPENDIX G—FACILITATOR SCRIPT

FACILITATOR SCRIPT

**GENERAL INSTRUCTIONS FOR
New York State Alternate Assessment
(NYSAA)
STANDARD SETTING GROUP FACILITATORS**

English Language Arts, GRADES 3-8 & HS

Mathematics, GRADES 3-8 & HS

Science, GRADES 4, 8 & HS

Social Studies, GRADE HS

June 11-12, 2014

Preliminaries

Introductions:

- Welcome group; introduce yourself (name, affiliation, a little background information).
- Have each participant introduce him/herself.
- Ask each participant to sign a nondisclosure form. Do not proceed until a signed nondisclosure form has been collected from each participant.

Become Familiar with the NYSAA

Overview: Some of the panelists administered the assessment to students, while others did not. In order to ensure that all panelists have an understanding of the knowledge and skills assessed, thoroughly review the student datafolios with the group, walking through the quick reference guide and the NYSAA rubric. You will have a full Administration Manual including a HS sample datafolio that a panelist can reference if he/she would like.

Activities:

1. Introduce the NYSAA materials:
 - a) Have panelists review the quick reference guide;
 - b) Have panelists review the practice datafolio set;
 - c) Have panelists review the scoring rubric; and
 - d) Have panelists review in pairs the NYSAA Frameworks.
2. Once they are done reviewing the materials, give panelists a few minutes to discuss any questions or issues that arose as they were reviewing the materials.

Discuss Alternate Performance Level Descriptors (APLDs)

Overview: In order to establish an understanding of the expected performance of students on the alternate assessment, panelists must have a clear understanding of:

- the definition of the four performance levels, and
- what the key characteristics are that distinguish students in the four performance level categories.

The purpose of this activity is for the panelists to come to consensus about what characterizes students in each performance level category. This activity is critical since the ratings panelists will be making will be based on these understandings.

Activities:

- Introduce the task. In this activity they will:
 - individually review the Alternate Performance Level Descriptors; and
 - discuss the descriptors as a group.
- Pass out the Alternate Performance Level Descriptors and have panelists review them individually. Panelists can make notes if they like. The goal here is for the panelists to come to a **common understanding** of what it means to be in each performance level. It is not unusual for panelists to disagree with the descriptors they will see; almost certainly there will be some panelists who will want to change them. However, the task at hand is for panelists to have a common understanding of what knowledge, skills, and understanding (KSUs) are described by each of the Alternate Performance Level Descriptors.
- During the discussion of each performance level, using chart paper, create a bulleted list for each level, specifying the specific knowledge, skills and understanding that best describe students in that level. The panelists want to answer the question: What knowledge, skills and understanding must a student demonstrate in order to be classified in the Level 1/Not Meeting Learning Standards? Or, put another way: What are the most important knowledge, skills and abilities that distinguish a Level 1(Not Meeting Learning Standards) student from a student in the Level 2 (Partially Meeting Learning Standards) category? They will then repeat this process for the Level 3 (Meeting Learning Standards), and Level 4 (Meeting Learning Standards with Distinction) categories. At this stage panelists might think of what a student at the threshold between categories might produce for work.
- Have the panelists identify the most important characteristics describing students at each performance level and record those as a bulleted list on chart paper. The list should be posted on the wall for panelists to refer to as they complete the three rounds of rating.

Practice Round

Overview of Practice Round: The primary purpose of the practice round is for panelists to become familiar with the rating task with three practice Bodies of Work (BOWs). Panelists will review the BOWs with the facilitator and discuss them as a group. The facilitator will walk the panelists through the practice ratings, engage the panelists in a readiness discussion and check for understanding. If any of the panelists indicate an incomplete understanding of the practice rating task, then the facilitator will continue to work with the panelists to clarify any misconceptions before proceeding to Round 1.

Activities:

1. Orient panelists to the set of datafolios. Point out that the datafolios are presented in order, from lowest scoring to highest scoring. Explain that the student datafolios represent how the typical student at each selected total score point performed.
2. The panelists will begin by reviewing the datafolios with the facilitator. As they are reviewing the datafolios, they should consider the knowledge, skills, and understanding demonstrated in each datafolio and how they relate to the performance level descriptors. The purpose of this step is to thoroughly familiarize panelists with the materials and with the rating process as well as to allow panelists to get an initial sense of how they feel the datafolios should be categorized.

Steps for walking through the practice BOWs:

- a) For the first datafolio walk through it as a group, read through Data Summary Sheets for each sample, Discuss the KSUs represented by this sample. Discuss the APLDs and determine which APLD best fits the KSUs represented in this profile. This profile will be selected to be clearly in the first performance level.
 - b) For the second datafolio walk through it as a group. Lead them through discussion of their rationale. Remind panelists that for a datafolio to fall into the next performance level they are looking for the profile to be just over the line between the first and second performance level-they are not looking for a solid second level performance.
 - c) For the third datafolio have the panelists rate it on his or her own and then lead them through a discussion of the rationale. Ask for a show of hands indicating which panelists placed the datafolio in each of the performance levels. Have the panelists discuss the rationale behind their placements.
3. Panelists may want to take notes as they work if there are particular points they would like to discuss with their colleagues.
 4. Make sure panelists know to enter their ratings on the practice rating form.
 5. Go over the practice rating form with panelists:
 - a) Have panelists write their ID number on the rating form. The ID number is on their name tags.
 - b) Lead panelists through a step-by-step demonstration of how to fill in the rating form.

Readiness Discussion

After the panelists have completed the practice rating form, lead a readiness discussion by posing the questions below.

The purpose of this discussion is to determine how well each panelist understands the rating task, to correct any misunderstandings, and if necessary, to identify panelists whose ratings should be excluded from the standard setting if their understanding doesn't improve.

The "correct" answers for each of the questions are listed directly under each question. Please guide the discussion to ensure each panelist understands the rating task. **Make sure any questions or concerns are resolved prior to moving on.**

- What type of information should I be considering when making my classification?
 - The PLDs.
 - The holistic set of knowledge, skills and understanding represented in each BOW.
 - My understanding of who these students are.

- Why is it important to refer to the APLDs?
 - The APLDs specify the knowledge, skills and understanding demonstrated by the student. It is these KSUs that need to correspond to the KSUs outlined in each performance level.

- If a student BOW reflects KSUs that "just barely" correspond to Proficient, what performance level should be assigned to the BOW?
 - Proficient: A BOW does not need to represent solid Proficient performance to be classified in that level, it just needs to make it into the category.

Round 1 Ratings

Overview of Round 1: The primary purpose of Round 1 is to ask the panelists to make their initial determination as to which performance level category each BOW should be classified into. In this round, panelists will be working individually, without discussion with their colleagues.

Activities:

- Make sure panelists have the following materials:
 - Rating form
 - Set of BOWs
 - Alternate Performance Level Descriptors
- Orient panelists to the set of BOWs. Point out that the BOWs are presented in order, from lowest scoring to highest. Make sure panelists understand that, even though the BOWs are presented from lowest- to highest-scoring, their own ratings do not need to be in strictly increasing order.
- Starting with the first BOW, the panelists will review each BOW in turn. As they are reviewing the BOWs, the panelists should keep in mind the Alternate Performance Level Descriptors. They should consider the knowledge, skills, and understanding demonstrated by each BOW and how they relate to the descriptors of the performance levels. The purpose of this step is for panelists to make their initial determinations as to how the BOWs should be categorized into the two performance levels. The panelists are free to make notes on the BOWs, sort them into piles, use sticky notes, or use whatever system helps them to keep track of their categorizations.
- Panelists may want to take notes as they work if there are particular points they would like to discuss with their colleagues in Round 2.
 - Have panelists write their ID and round number on the rating form. The ID number is on their name tags.
 - Briefly remind them how to fill in the rating form.
 - Answer questions the panelists may have about the work in Round 1.
 - Once everyone understands what they are to do in Round 1, tell them to begin.
- As panelists complete the task, ask them to carefully inspect their rating forms to ensure they are filled out properly.
 - The ID and round number must be filled in.
 - Each BOW must be assigned to one and only one performance level.
 - Reiterate that although the BOWs are presented in order from lowest- to highest-scoring, the panelists' category assignments do not need to be in strictly increasing order.
- Facilitators should bring all the completed rating forms together to the data analysis work room for tabulation. Prior to submitting them, however, using a show of hands, indicate on a piece of chart paper how many panelists assigned each BOW to each performance level category. This chart will be used for the Round 2 discussions.

Tabulation of Round 1 Results

Tabulation of Round 1 results will be completed as quickly as possible after receipt of the rating forms. While the tabulation occurs, the group may take a break.

Round 2 Ratings

Overview of Round 2: In Round 2, the panelists will have an opportunity to discuss their Round 1 placements and to revise their ratings on the basis of that discussion. Prior to beginning the Round 2 discussions, the psychometrician will share the group's average cut point based on the Round 1 ratings and present impact data.

Focusing on any BOWs for which there was disagreement as to how they should be categorized, the panelists will discuss why they categorized each BOW as they did, making sure that all different points of view are included in the discussion.

Once panelists have reviewed and discussed the Round 1 categorizations, they will make their Round 2 ratings.

Activities:

- Make sure panelists have the following materials:
 - Rating form
 - Set of BOWs
 - Alternate Performance Level Descriptors
- Have panelists write their ID number and Round 2 on the rating form.
- Provide an overview of Round 2. Paraphrase the following:
 - As in Round 1, the primary purpose is to categorize each BOW into the performance level category where you believe it belongs.
 - Each panelist needs to base his/her judgments on his/her experience with the content area, understanding of students, the definition of each performance level category, discussions with other panelists, and the knowledge, skills, and understanding demonstrated in the datafolio.
- Show the panelists how the BOWs would be categorized based on the room average Round 1 cut point placement.
- Remind panelists that they will be discussing the BOWs with their colleagues, but that they will be categorizing the BOWs individually. It is not necessary for the panelists to reach consensus about how to categorize each BOW.
- Give panelists an opportunity to ask questions about the task for Round 2.
- Beginning with the first BOW for which there is disagreement as to how it should be categorized, the panelists should begin discussing the categorization of the BOWs according to the Round 1 ratings.
 - Panelists only need to discuss those BOWs for which there was disagreement as to how they should be categorized.
 - Starting with the first BOW for which there was disagreement as to the categorization, the group will discuss the KSUs demonstrated in each BOW and how those KSUs align to the APLDs.
 - Panelists should be encouraged to listen to their colleagues as well as express their own points of view.
 - If the panelists hear a logic/rationale/argument that they did not consider and that they feel is compelling, then they may adjust their ratings to incorporate that information.
 - On the basis of the discussions and the feedback presented, panelists should make their Round 2 ratings.
 - The group does not have to achieve consensus. If panelists honestly disagree, that is fine. We are trying to get the best judgment of each panelist. Panelists should not feel compelled or coerced into making a rating with which they disagree.

Encourage the panelists to use the discussion and feedback to assess how stringent or lenient a judge they are. If a panelist is categorizing BOWs consistently higher or lower than the group, he or she may have a different understanding of the Alternate Performance Level Descriptors than the rest of the group. **It is acceptable for panelists to disagree, but that disagreement should be based on a common understanding of the Alternate Performance Level Descriptors.**

- When the group has completed their Round 2 ratings, collect the rating forms. When you collect the rating forms, carefully inspect them to ensure they are filled out properly.
 - The ID and round number must be filled in.
 - Each BOW must have one (and only one) rating.
- Facilitators should bring all the completed rating forms together to the data analysis work room for tabulation.

Tabulation of Round 2 Results

Tabulation of Round 2 results will be completed as quickly as possible after receipt of the rating forms. While the tabulation occurs, the group may take a break.

Round 3 Ratings

Overview of Round 3: In Round 3, the panelists will have a final opportunity to discuss their Round 2 placements and to revise their ratings on the basis of that discussion. Prior to beginning the Round 3 discussions, the psychometrician will share the group's average cut point based on the Round 2 ratings and impact data. The impact data that will be shared here will be current impact data as well as historical impact data. In this round, panelists will focus their discussion specifically on the 5 BOWs above and below the group average cut.

Once panelists have reviewed and discussed the Round 2 categorizations, they will make their final ratings.

Activities:

- Make sure panelists have the following materials:
 - Rating form
 - Set of BOWs
 - Alternate Performance Level Descriptors
- Have panelists write their ID number and Round 3 on the rating form.
- Provide an overview of Round 3. Paraphrase the following:
 - As in Round 2, the primary purpose is to categorize each BOW into the performance level category where you believe it belongs.
 - Each panelist needs to base his/her judgments on his/her experience with the content area, understanding of students, the definition of each performance level category, discussions with other panelists, and the knowledge, skills, and understanding demonstrated in the datafolio.
- Review the feedback information with the panelists.
 - Show the panelists how the BOWs would be categorized based on the room average Round 2 cut point placement.
- Remind panelists that they will be discussing the BOWs with their colleagues, but that they will be categorizing the BOWs individually. It is not necessary for the panelists to reach consensus about how to categorize each BOW.
- Give panelists an opportunity to ask questions about the feedback information or about the task for Round 3.
- Beginning with the first cut point, the panelists will discuss their Round 2 ratings.
 - The discussion should focus on the 5 BOWs below and above the cut point based on the group average cuts.
 - Starting with the fifth BOW below the cut, the group will discuss the KSUs demonstrated in each BOW and how those KSUs align to the APLDs.
 - Panelists should be encouraged to listen to their colleagues as well as express their own points of view.
 - If the panelists hear a logic/rationale/argument that they did not consider and that they feel is compelling, then they may adjust their ratings to incorporate that information.
 - On the basis of the discussions and the feedback presented, panelists should make their final ratings.
 - The group does not have to achieve consensus. If panelists honestly disagree, that is fine. We are trying to get the best judgment of each panelist. Panelists should not feel compelled or coerced into making a rating with which they disagree.

Encourage the panelists to use the discussion and feedback to assess how stringent or lenient a judge they are. If a panelist is categorizing BOWs consistently higher or lower than the group, he or she may have a different understanding of the Alternate Performance Level Descriptors

than the rest of the group. It is acceptable for panelists to disagree, but that disagreement should be based on a common understanding of the Alternate Performance Level Descriptors.

- When the group has completed their final ratings, collect the rating forms. When you collect the rating forms, carefully inspect them to ensure they are filled out properly.
 - The ID and round number must be filled in.
 - Each BOW for Round 3 must have one (and only one) rating.
 - Return the completed rating forms to the data analysis work room.

Additional Grades

After the completion of the first grade, the process will repeat for each subsequent grade. For subsequent grades, all steps except the training will be repeated:

Discuss Alternate Performance Level Descriptors (PLDs)

Overview: In order to establish an understanding of the expected performance of students on the test, panelists must have a clear understanding of:

- the definition of the four performance levels, and
- what the key characteristics are that distinguish students in the four performance level categories.

The purpose of this activity is for the panelists to come to consensus about what characterizes students in each performance level category. This activity is critical since the ratings panelists will be making will be based on these understandings.

Activities:

- Introduce the task. In this activity they will:
 - Individually review the Alternate Performance Level Descriptors; and
 - discuss the descriptors as a group.
- Pass out the Alternate Performance Level Descriptors and have panelists review them individually. Panelists can make notes if they like. The goal here is for the panelists to come to a **common** understanding of what it means to be in each performance level. It is not unusual for panelists to disagree with the descriptors they will see; almost certainly there will be some panelists who will want to change them. However, the task at hand is for panelists to have a common understanding of what knowledge, skills, and understanding (KSUs) are described by each of the Alternate Performance Level Descriptors.
- During the discussion of each performance level, using chart paper, create a bulleted list for each level, specifying the knowledge, skills and understanding that best describe students in that level. The panelists want to answer the question: What knowledge, skills and understanding must a student demonstrate in order to be classified in the Level 1/Not Meeting Learning Standards? Or, put another way: What are the most important knowledge, skills and abilities that distinguish a Level 1(Not Meeting Learning Standards) student from a student in the Level 2 (Partially Meeting Learning Standards) category? They will then repeat this process for the Level 3 (Meeting Learning Standards), and Level 4 (Meeting Learning Standards with Distinction) categories. At this stage panelists might think of what a student at the threshold between categories might produce for work.

- Have the panelists identify the most important characteristics describing students at each performance level and record those as a bulleted list on chart paper. The list should be posted on the wall for panelists to refer to as they complete the three rounds of rating.

Familiarize Panelists with Datafolios

In this activity, the panelists will review and become familiar with the new set of datafolios that will be categorized into performance levels.

Activities:

1. Orient panelists to the set of datafolios. Remind panelists that the datafolios are presented in order, from lowest scoring to highest scoring.
2. Allow for time for panelists to discuss any issues or ask any questions that arise during the review of the datafolios.

Complete the 3 Rounds of Ratings in the same way as above.

Complete Final Evaluation Form

After the rounds of ratings are completed, have the panelists fill out the final evaluation form. Emphasize that their honest feedback is important.

Additional Information

In addition to the process of standard setting there are two additional issues of which you should be aware. They will not necessarily come up in the standard setting process, but is information you should be aware of.

Level of Complexity

The NYSAA Extensions, built on the three complexity levels, will be weighted. The most complex Extensions will have a heavier weight than the middle level Extensions and the lowest level Extensions will not have an additive constant applied.

Independence documentation

The independence documentation on the Data Summary Sheet is not part of the student's overall performance level. It is documented by teachers to use for their own understanding of student performance.

Baseline and Final data documentation

The Baseline documentation on the Data Summary Sheet is not part of the student's overall performance level. It is documented by teachers to demonstrate the student is being assessed at an appropriate complexity level. Only the Final documentation on the Data Summary Sheet is factored in to the student's overall performance level.