Common Formative Assessment

Performance Assessment Items

Handout Packet
Here are seven general characteristics of performance tasks:

Performance tasks call for the application of knowledge and skills, not just recall or recognition.

In other words, the learner must actually use their learning to perform. These tasks typically yield a tangible product (e.g., graphic display, blog post) or performance (e.g., oral presentation, debate) that serve as evidence of their understanding and proficiency.

Performance tasks are open-ended and typically do not yield a single, correct answer.

Unlike selected- or brief constructed-response items that seek a “right” answer, performance tasks are open-ended. Thus, there can be different responses to the task that still meet success criteria. These tasks are also open in terms of process; i.e., there is typically not a single way of accomplishing the task.

Performance tasks establish novel and authentic contexts for performance.

These tasks present realistic conditions and constraints for students to navigate. For example, a mathematics task would present students with a never-before-seen problem that cannot be solved by simply “plugging in” numbers into a memorized algorithm. In an authentic task, students need to consider goals, audience, obstacles, and options to achieve a successful product or performance. Authentic tasks have a side benefit—they convey purpose and relevance to students, helping learners see a reason for putting forth effort in preparing for them.

Performance tasks provide evidence of understanding via transfer.

Understanding is revealed when students can transfer their learning to new and “messy” situations. Note that not all performances require transfer. For example, playing a musical instrument by following the notes or conducting a step-by-step science lab require minimal transfer. In contrast, rich performance tasks are open-ended and call “higher-order thinking” and the thoughtful application of knowledge and skills in context, rather than a scripted or formulaic performance.

Performance tasks are multi-faceted.

Unlike traditional test “items” that typically assess a single skill or fact, performance tasks are more complex. They involve multiple steps and thus can be used to assess several standards or outcomes.

Performance tasks can integrate two or more subjects as well as 21st century skills.
In the wider world beyond the school, most issues and problems do not present themselves neatly within subject area “silos.” While performance tasks can certainly be content-specific (e.g., mathematics, science, social studies), they also provide a vehicle for integrating two or more subjects and/or weaving in 21st century skills and Habits of Mind. One natural way of integrating subjects is to include a reading, research, and/or communication component (e.g., writing, graphics, oral or technology presentation) to tasks in content areas like social studies, science, health, business, health/physical education. Such tasks encourage students to see meaningful learning as integrated, rather than something that occurs in isolated subjects and segments.

Performances on open-ended tasks are evaluated with established criteria and rubrics.

Since these tasks do not yield a single answer, student products and performances should be judged against appropriate criteria aligned to the goals being assessed. Clearly defined and aligned criteria enable defensible, judgment-based evaluation. More detailed scoring rubrics, based on criteria, are used to profile varying levels of understanding and proficiency.

For more information about performance tasks, see Jay McTighe’s blog.

What?...So What?...Now What?

What?
List key ideas & interpretations

So What?
List what you learned from

Now What?
List the implications & applications for your classroom/building.
Formative Performance Assessment (FPA) Characteristics

(McTighe, 2015)
Formative Performance Assessment Characteristics...

- call for the application of knowledge and skills, not just recall or recognition.
- are open-ended and typically do not yield a single, correct answer.
- establish novel and authentic contexts for performance.
- provide evidence of understanding via transfer.
- are multi-faceted.
- can integrate two or more subjects as well as 21st century skills.
- are evaluated with established criteria and rubrics.

## Formative Instructional and Assessment Tasks

### Carter’s Candy Company

#### 5.MD.3-Task 1

<table>
<thead>
<tr>
<th>Domain</th>
<th>Measurement and Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster</td>
<td>Geometric measurement: understand concepts of volume and relate volume to multiplication and to addition.</td>
</tr>
</tbody>
</table>
| Standard(s)  | 5.MD.3 Recognize volume as an attribute of solid figures and understand concepts of volume measurement.  
|              | a. A cube with side length 1 unit, called a “unit cube” is said to have “one cubic unit” of volume, and can be used to measure volume.  
|              | b. A solid figure which can be packed without gaps or overlaps using \( n \) unit cubes is said to have a volume of \( n \) cubic units. |
| Materials    | Paper and pencil, unit cubes, isometric graph paper                                    |
| Task         | Carter’s Candy Company is selling a new type of chocolate. They have decided to sell the candy in packages of 24. You are leading a team in charge of developing a box for the candy.  
|              | Use unit cubes to build all of the possible boxes for the package of candy. How many possibilities are there? Record the dimensions and volume of each box. What do you notice about all of the volumes?  
|              | After determining all of the possible boxes, you must make a recommendation to the president of the company about which box should be used. Write a paragraph explaining which box would be best. Make sure to explain your reasons for choosing this box.  
|              | Two of your team members get in an argument about the boxes. Cathy says that a 1x24 box is the same as a 24x1 box. Curtis says that these dimensions would lead to two different boxes. Who do you agree with? Why? |

### Formative Instructional and Assessment Tasks

<table>
<thead>
<tr>
<th>Rubric</th>
<th>Level I</th>
<th>Level II</th>
<th>Level III</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Limited Performance</strong></td>
<td>• Student is able to build some boxes with a volume of 24.</td>
<td>• Student is able to build boxes that have a volume of 24 but may not find all of the possible combinations.</td>
<td>• Student recognizes that all the boxes will have a volume of 24, since that is how many pieces of candy it must contain.</td>
</tr>
<tr>
<td></td>
<td>• Student is unable to explain whether Cathy or Curtis is correct.</td>
<td>• Student recognizes the connection between volume and dimensions.</td>
<td>• Student identifies all the combinations that will lead to a volume of 24: 1x24, 2x12, 3x8, 4x6 (and the reverse)</td>
</tr>
<tr>
<td></td>
<td>• Student may recommend a box but it unable to use math language to justify reasoning.</td>
<td>• Student recommends which box to use but explanation lacks detail or is unclear.</td>
<td>• Student identifies relationship between volume and linear dimensions.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Student is unsure how to settle the argument between Cathy and Curtis.</td>
<td>• Student chooses a side – Cathy or Curtis. Student demonstrates that both boxes have the same volume and the same dimensions. However, because of packaging on the outside of the box, and depending on how the candy is shaped and needs to be stacked, these could be considered different boxes. Allow a variety of responses, as long as student demonstrates understanding of the concept of volume.</td>
</tr>
</tbody>
</table>

**Standards for Mathematical Practice**

1. Makes sense and perseveres in solving problems.
2. Reasons abstractly and quantitatively.
3. Constructs viable arguments and critiques the reasoning of others.
5. Uses appropriate tools strategically.
6. Attends to precision.
7. Looks for and makes use of structure.
8. Looks for and expresses regularity in repeated reasoning.
Formative Instructional and Assessment Tasks
Carter’s Candy Company

Carter’s Candy Company is selling a new type of chocolate. They have decided to sell the candy in packages of 24. You are leading a team in charge of developing a box for the candy.

A. Use unit cubes to build all of the possible boxes for the package of candy. How many possibilities are there? Record the dimensions and volume of each box. What do you notice about all of the volumes?

B. After determining all of the possible boxes, you must make a recommendation to the president of the company about which box should be used. Write a paragraph explaining which box would be best. Make sure to explain your reasons for choosing this box.

C. Two of your team members get in an argument about the boxes. Cathy says that a 1x24 box is the same as a 24x1 box. Curtis says that these dimensions would lead to two different boxes. Who do you agree with? Why?
ELA II Writing Performance Assessment with Rubric

Many people enjoy one season of the year more than the others. Choose the season you enjoy the most. Write an explanatory essay for your teacher expressing clear details about that season, and give reasons why you most enjoy this time. Include facts, details, and precise language as they are appropriate in your writing.

<table>
<thead>
<tr>
<th>Score</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
<th>NS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Organization/Purpose</strong></td>
<td>The response has a clear and effective organizational structure, creating a sense of unity and completeness. The response is fully sustained, and consistently and purposefully focused:</td>
<td>The response has an evident organizational structure and a sense of completeness, though there may be minor flaws and some ideas may be loosely connected. The response is adequately sustained and generally focused:</td>
<td>The response has an inconsistent organizational structure, and flaws are evident. The response is somewhat sustained and may have a minor drift in focus:</td>
<td>The response has little or no discernible organizational structure. The response may be related to the topic but may provide little or no focus:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• controlling or main idea of a topic is clearly communicated, and the focus is strongly maintained for the purpose, audience, and task</td>
<td>• controlling or main idea of a topic is clear, and the focus is mostly maintained for the purpose, audience, and task</td>
<td>• controlling or main idea of a topic may be somewhat unclear, or the focus may be insufficiently sustained for the purpose, audience, and task</td>
<td>• controlling or main idea may be confusing or ambiguous; response may be too brief or the focus may drift from the purpose, audience, or task</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• consistent use of a variety of transitional strategies to clarify the relationships between and among ideas</td>
<td>• adequate use of transitional strategies with some variety to clarify the relationships between and among ideas</td>
<td>• inconsistent use of transitional strategies and/or little variety</td>
<td>• few or no transitional strategies are evident</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• effective introduction and conclusion</td>
<td>• adequate introduction and conclusion</td>
<td>• introduction or conclusion, if present, may be weak</td>
<td>• introduction and/or conclusion may be missing</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• logical progression of ideas from beginning to end; strong connections between and among ideas with some syntactic variety</td>
<td>• adequate progression of ideas from beginning to end; adequate connections between and among ideas</td>
<td>• uneven progression of ideas from beginning to end; and/or formulaic; inconsistent or unclear connections between and among ideas</td>
<td>• frequent extraneous ideas may be evident</td>
<td></td>
</tr>
</tbody>
</table>

### 4-Point Explanatory Writing Rubric

<table>
<thead>
<tr>
<th>Score</th>
<th>Evidence/Elaboration</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
<th>NS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>The response provides thorough and convincing support/evidence for the controlling idea and supporting idea(s) that include the effective use of facts and details. The response clearly and effectively elaborates ideas, using precise language:</td>
<td>The response provides adequate support/evidence for the controlling idea and supporting idea(s) that include facts and details. The response adequately elaborates ideas, employing a mix of precise and more general language:</td>
<td>The response provides uneven, cursory support/evidence for the controlling idea and supporting idea(s) that include facts and details. The response elaborates ideas unevenly, using simplistic language:</td>
<td>The response provides minimal support/evidence for the controlling idea and supporting idea(s) that include facts and details. The response is vague, confusing, or lacks clarity:</td>
<td>The response provides minimal support/evidence for the controlling idea and supporting idea(s) that include facts and details. The response is vague, confusing, or lacks clarity:</td>
<td>Insufficient (includes copied text) Explicit refusal Statement of non-understanding Predomately in another language</td>
</tr>
<tr>
<td></td>
<td>• comprehensive support is integrated, relevant, and specific</td>
<td>• adequate support is integrated and relevant, yet may be general</td>
<td>• some support may be weakly integrated, imprecise, repetitive vague, and/or copied</td>
<td>• support is minimal, irrelevant, absent, incorrectly used, or predominately copied</td>
<td>• support is minimal, irrelevant, absent, incorrectly used, or predominately copied</td>
<td>• Insufficient (includes copied text) Explicit refusal Statement of non-understanding Predomately in another language</td>
</tr>
<tr>
<td></td>
<td>• effective use of a variety of elaborative techniques*</td>
<td>• adequate use of some elaborative techniques*</td>
<td>• weak or uneven use of elaborative techniques*</td>
<td>• minimal, if any, use of elaborative techniques*</td>
<td>• minimal, if any, use of elaborative techniques*</td>
<td>• Insufficient (includes copied text) Explicit refusal Statement of non-understanding Predomately in another language</td>
</tr>
<tr>
<td></td>
<td>• vocabulary is clearly appropriate for the audience and purpose</td>
<td>• vocabulary is generally appropriate for the audience and purpose</td>
<td>• vocabulary use is uneven or somewhat ineffective for the audience and purpose</td>
<td>• vocabulary is limited or ineffective for the audience and purpose</td>
<td>• vocabulary is limited or ineffective for the audience and purpose</td>
<td>• Insufficient (includes copied text) Explicit refusal Statement of non-understanding Predomately in another language</td>
</tr>
<tr>
<td></td>
<td>• effective, appropriate style enhances content</td>
<td>• generally appropriate style is evident</td>
<td>• inconsistent or weak attempt to create appropriate style</td>
<td>• little or no evidence of appropriate style</td>
<td>• little or no evidence of appropriate style</td>
<td>• Insufficient (includes copied text) Explicit refusal Statement of non-understanding Predomately in another language</td>
</tr>
</tbody>
</table>

*Elaborative techniques may include the use of personal experiences that support the controlling idea.

## 2-Point Explanatory Writing Rubric

<table>
<thead>
<tr>
<th>Score</th>
<th>2</th>
<th>1</th>
<th>0</th>
<th>NS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Conventions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The response demonstrates an adequate command of conventions:</td>
<td>The response demonstrates a partial command of conventions:</td>
<td>The response demonstrates little or no command of conventions:</td>
<td>Insufficient (includes copied text)</td>
<td></td>
</tr>
<tr>
<td>• adequate use of correct sentence formation, punctuation, capitalization, grammar usage, and spelling</td>
<td>• limited use of correct sentence formation, punctuation, capitalization, grammar usage, and spelling</td>
<td>• infrequent use of correct sentence formation, punctuation, capitalization, grammar usage, and spelling</td>
<td>Explicit refusal</td>
<td></td>
</tr>
</tbody>
</table>

**Holistic Scoring:**

- **Variety:** A range of errors includes formation, punctuation, capitalization, grammar usage, and spelling.
- **Severity:** Basic errors are more heavily weighted than higher-level errors.
- **Density:** The proportion of errors to the amount of writing done well. This includes the ratio of errors to the length of the piece.

---

Developing a Formative Performance Assessment

1. Identify Priority/Power/Essential Learning Standards for a unit of study.

2. Determine standards that relate to the Priority Standards – including interdisciplinary standards.

3. “Unwrap” the chosen standards to determine concepts and skills.

4. “Deconstruct” the “unwrapped” standards to determine the criteria to be assessed.

5. Determine levels of performance for each criterion.

6. Determine the appropriate performance types that will address the criteria.

7. Develop a rubric that assesses the chosen criteria.

8. Share rubric with students prior to instruction.

Students may have input on the creation of classroom “student friendly” rubric.


Evaluating a Formative Performance Assessment

1. “Activity requires students to construct a response, create a product, or perform a demonstration.

2. Open-ended---may not always yield a single correct answer or solution method.

3. Evaluations of student products or performances are based on scoring rubrics provided to students in advance of performance.

4. “Highly engaging for students; connects or applies content knowledge and skills to real-world situations.

5. Promotes critical thinking – students must ‘show what they know’ through the use of higher-level thinking skills.

6. Student responses provide credible evidence that standards have or have not been met.

7. Motivates all students to be proficient.

8. “Utilizes collaborative learning process but with individual accountability.


10. Offers multiple opportunities for students to revise work using scoring guide feedback.

11. Traditional tests used as ‘concurrent validity’ measure – together with performance assessments, they provide ‘multiple measures’ of student achievement.”

## Desired Characteristics of Criteria for Classroom Rubrics

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>The criteria are…</td>
<td><strong>Appropriate</strong> Each criterion represents an aspect of a standard, curricular goal, or instructional goal or objective that students are intended to learn.</td>
</tr>
<tr>
<td></td>
<td><strong>Definable</strong> Each criterion has a clear, agreed-upon meaning that both students and teachers understand.</td>
</tr>
<tr>
<td></td>
<td><strong>Observable</strong> Each criterion describes a quality in the performance that can be perceived (seen or heard, usually) by someone other than the person performing.</td>
</tr>
<tr>
<td></td>
<td><strong>Distinct from one another</strong> Each criterion identifies a separate aspect of the learning outcomes the performance is intended to assess.</td>
</tr>
<tr>
<td></td>
<td><strong>Complete</strong> All the criteria together describe the whole of the learning outcomes the performance is intended to assess.</td>
</tr>
<tr>
<td></td>
<td><strong>Able to support descriptions along a continuum of quality</strong> Each criterion can be described over a range of performance levels.</td>
</tr>
</tbody>
</table>

Developing FPA Rubrics

1. Determine criteria that demonstrate the learning targets.
2. Decide the continuum of levels.
3. Write a description of performance quality for each level of each criterion.
Developing Formative Performance Assessment Rubrics

1. Determine criteria that demonstrate the learning targets.
2. Decide the continuum of levels.
3. Write continuum from highest to lowest levels with highest on the left.
4. Write a description of performance quality for each level of each criterion.
5. Determine labels for each level.

Evaluating Formative Performance Assessment Rubrics

“The biggest mistake teachers make when they use rubrics with performance assessment is that they focus on the task, the product, and not the learning outcome or proficiency the task is supposed to get students to demonstrate.”

Avoid the following flaws when creating a rubric:

1. Confounding the outcomes being measured [scoring more than one content-area skill at a time, without recognizing them as separate skills]
2. Scoring for extraneous features (e.g., neatness, color, etc.)
3. Scoring by counting up parts or components rather than by looking for evidence of proficiency in the outcome(s) being measured
4. Scoring for things student have not been cued to do
5. Scoring products rather than outcomes

## Math Problem-Solving

<table>
<thead>
<tr>
<th>Your score</th>
<th>SHOWING MATH KNOWLEDGE (Can you do the problem correctly?)</th>
<th>USING PROBLEM-SOLVING STRATEGIES (How do you solve the problem?)</th>
<th>WRITING AN EXPLANATION (Can you explain your work?)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>• I figure out the <strong>correct answer</strong>.</td>
<td>• I use <strong>all the important information</strong> from the problem.</td>
<td>• I write <strong>what</strong> I did and <strong>why</strong> I did it.</td>
</tr>
<tr>
<td></td>
<td>• I solve the problem with no mistakes.</td>
<td>• I show <strong>all the steps</strong> I used to solve the problem.</td>
<td>• I explain each step of my work.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• I make a <strong>drawing</strong> to show how I solved the problem.</td>
<td>• I use <strong>math words and strategy names</strong>.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• I write the <strong>answer in a complete sentence</strong> at the end of my explanation.</td>
</tr>
<tr>
<td>4</td>
<td>• I figure out the <strong>correct answer</strong>.</td>
<td>• I use <strong>most of the important information</strong> from the problem.</td>
<td>• I write <strong>what</strong> I did and a little about <strong>why</strong> I did it.</td>
</tr>
<tr>
<td></td>
<td>• I solve the problem, but I make a few small mistakes.</td>
<td>• I show <strong>most of the steps</strong> I used to solve the problem.</td>
<td>• I explain <strong>most</strong> of my work.</td>
</tr>
<tr>
<td>3</td>
<td>• I figure out <strong>part of the answer.</strong></td>
<td>• I use <strong>some of the important information</strong> from the problem.</td>
<td>• I write a little about <strong>what</strong> I did or <strong>why</strong> I did it, but not both.</td>
</tr>
<tr>
<td></td>
<td>• I try to solve the problem, but I make some big mistakes.</td>
<td>• I show <strong>some of the steps</strong> I used to solve the problem.</td>
<td>• I explain <strong>some</strong> of my work.</td>
</tr>
<tr>
<td>2</td>
<td>• I try to solve the problem, but I don’t understand it.</td>
<td>• I use <strong>very little important information</strong> from the problem.</td>
<td>• I write something that doesn’t make sense.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• I show <strong>almost none of the steps</strong> I used to solve the problem.</td>
<td>• I write an <strong>unclear answer</strong>.</td>
</tr>
<tr>
<td>1</td>
<td>• I <strong>don’t try</strong> to solve the problem.</td>
<td>• I show <strong>no steps</strong> that I used to solve the problem.</td>
<td>• I <strong>don’t write anything</strong> to explain how I solved the problem.</td>
</tr>
</tbody>
</table>

Choosing Anchor Papers

1. Separate student responses into two stacks: proficient and not proficient.
2. Using your scoring rubric, separate the proficient papers in two levels of performance: exemplar and proficient (or other terms you have used).
3. Check to see whether descriptions of the responses match the descriptors of your rubric.
4. Using your scoring rubric, separate the not proficient papers in two levels of performance: close and not yet (or other terms you have used).
5. Check to see whether descriptions of the responses match the descriptors of your rubric.
6. Score student papers using the rubric.
7. After determining samples for each score point, choose one that is a solid representative sample for each score point.
8. Annotate each anchor by writing a description of the criteria the response exhibits for that score point.

If the scoring process doesn’t work well, check to see that your rubric has definitive levels. Re-evaluate your prompt and/or rubric.