

# *Developing and Using Performance Assessments*

Presented by

Christopher R. Gareis, Ed.D.

Professor of Education  
The College of William and Mary  
Williamsburg, Virginia USA  
crgare@wm.edu

Some materials in this packet are adapted from:

Gareis, C. R., & Grant, L. W. (2015). *Teacher-Made Assessments: How to Connect Curriculum, Instruction, and Student Learning*. (2<sup>nd</sup> edition.) Routledge Press.

## Types of Performance Assessments

Performance assessments are characterized by their potential to prompt students to engage in higher-order thinking, undertake complex processes and skills, design and create original products, and apply deep subject-area understandings in ways that are authentic to the real-world and to the discipline. However, performance assessments can vary greatly in their uses and forms. The table below, adapted from training the Stanford Center for Assessment, Learning, & Equity (SCALE), defines several of the distinguishing characteristics and forms of performance assessments.

CHARACTERISTICS	TYPES			
	CONSTRUCTED-RESPONSE	STAND-ALONE	CURRICULUM-EMBEDDED	COMPLEX PROJECT
Number of Intended Learning Outcomes (ILOs)	1 – 2 ILOs	Multiple, subject-specific ILOs	A cogent set of subject-specific ILOs	A complex, integrative set of ILOs & broad aims
Level of Instructional Support from Teacher During Administration	Limited to clarification	Limited clarification & facilitation	Integrated instruction, facilitation, & feedback	Integrated instruction, facilitation, feedback, & guidance
Prescriptiveness of Student Response (i.e., Degree of Student Choice)	Fixed/ Convergent (typically little choice)	Convergent (limited choice)	Moderately Divergent (elements of choice in content and/or format of response)	Divergent (typically multiple opportunities for student choice)
Approximate Duration	A portion of a class period (≤ 60 minutes)	1 – 2 class periods (> 60 minutes)	Multiple class periods or days	Multiple weeks or a term

## A Protocol for Using Anchors to Establish Interrater Reliability with a Validated Rubric

Adapted by Jeri Thompson, Center for Assessment (2013) from *Quality Performance Assessment: A Guide for Schools and Districts* (2012). (Permission to reproduce and use is given when authorship is fully cited.)

**Purpose:** To establish interrater reliability for a performance assessment.

*Interrater reliability* is “the consistency with which two or more scorers apply grading criteria of an assessment, thereby resulting in stable assessment results among students, uninfluenced by factors that are not the intended criteria of learning” (Gareis & Grant, 2015, p. 183).

**Time:** 30 minutes to several hours. Time varies by the complexity of the task (e.g., number of intended learning outcomes being assessed, the type of student response format, the number of scorers among whom interrater reliability is being established, and the depth of subject-area expertise required of the scorers).

**Group Size:** 3 – 8

**Materials Needed:**

- Task prompt
- Teacher directions (if available)
- Rubric or other scoring instrument (e.g., checklist, rating scale)
- 1 extra rubric or scoring instrument for the recorder, who will tally scores for the group
- Scoring guidelines (if available)
- 2 or more anchors per gradation of performance

**Roles:** Choose a facilitator, timekeeper, and recorder.

**Process:**

1. **Norms:** The facilitator review the protocol process and norms with the group. (2 minutes)
2. **Examination:** Group members silently examine the task prompt, rubric, and scoring guidelines (if available).
3. **Clarifying questions:** Group members ask clarifying questions they have about the materials and process.
4. **Optional Step to Establish Criteria:** Facilitator provides a set of anchor papers, each pegged to one of the gradations on the rubric but not yet known to the group. Group members individually review and score this initial set of anchor papers. The whole group then shares ratings and works to achieve a consensus understanding of the gradations within the rubric.

5. **Read and score:** Using the rubric (or scoring instrument), group members independently and silently read the sampling of student work (i.e., *anchors*), ranking them as high, average, or low based on their overall impression. Group members then individually score and student work using the rubric. Scorers should note words and phrases in the rubric's performance level descriptors that best describe the qualities of the work and make notes to explain and justify their scores. It is important to note that there won't always be an example of every score point within a given set of student work. The scoring rubric and evidence in the student work should always be the basis for the score, rather than the *relative* strength or weakness of a piece. The student work sample must be truly aligned to the description of the assigned score for the integrity of the exercise to be preserved.
6. **Score sharing:** One at a time, team members share their score for each of the rubric categories – without explanation – as the recorder completes the group's score sheet.
7. **Discussion:**
  - a. The facilitator invites the group to consider where the differences in the scores occurred and why people scored differently for each rubric area – particularly the highest and lowest scores.
  - b. Group members explain and justify scores by pointing to specific language in the rubric and evidence in the student work.
  - c. Group members discuss each piece of student work, resolving issues centered on either the meaning of the rubric or the merit and validity of the evidence in the student work until consensus is reached.
8. **Determine interrater reliability:** Led by the facilitator, the group gauges the degree of interrater reliability (typically as high, moderate, or low, although this can also be gauged statistically with enough samples).
  - a. What did we notice about scoring student work and using the rubric?
  - b. How confident are we in the consistency of our use of the rubric?
  - c. What are the implications for the future use of this rubric?
9. **Optional debrief:** The group may discuss the following questions after the above process:
  - a. What revisions should be made to the task and instructions?
  - b. What would be the next steps for instructing a given student or group of students from the sample of work?
  - c. What are the implications for our curriculum or our instructional practices for future students?

**Directions:** Use the painting to answer the question below.



**Title:** *The First Thanksgiving 1621*  
**By:** J.L.G. Ferris  
**Date:** 1932

**Question:** The painting *The First Thanksgiving 1621* helps historians understand the relationship between the Wampanoag Indians and the Pilgrim settlers in 1621.

Do you agree or disagree? (Circle one.)

Briefly support your answer:

---

---

---

---

---

---

---

TASK-SPECIFIC  
RUBRIC

*The First Thanksgiving*

Constructed-Response Performance Assessment

RUBRIC

To answer the question correctly, students must notice both the date of the event (1621) and the date of the painting (1932) and understand that the intervening time makes it a less useful source for a historian trying to learn about the first Thanksgiving.

<b>Basic</b>	<b>Emergent</b>	<b>Proficient</b>
Student does not recognize the gap in time and instead takes the painting at face value or provides an irrelevant response.	Student evaluates reliability of source but does not fully explain the problems caused by the gap in time.	Student explains why the time gap limits the reliability of the source as evidence of what happened at the first Thanksgiving.

Rubric adapted from Stanford History Education Group "The First Thanksgiving" available at <https://sheg.stanford.edu/history-assessments>.

The painting *The First Thanksgiving of 1621* helps historians understand the relationship between the Wampanoag Indians and the Pilgrim settlers in 1621. Do you agree or disagree? Briefly support your answer.

Then, in a one-paragraph response, explain how the historical context in which the painting was created may be reflected in the interpretation of the event depicted in the painting itself. Your response should include accurate historical detail to support your claim.

Be sure to make use of good writing conventions (such as deliberate word choice, precise information, grammar, and mechanics) in order to clearly convey your thinking.

**INTEGRATED,  
TASK-SPECIFIC  
RUBRIC**

	<b>Not Evident (0)</b>	<b>Basic (1)</b>	<b>Emergent (2)</b>	<b>Proficient (3)</b>
<b>Determine the relevance of a source</b>	<input type="checkbox"/> Irrelevant response or no response provided	<input type="checkbox"/> Does not recognize the gap in time and instead takes the painting at face value	<input type="checkbox"/> Evaluates reliability of source but does not fully explain the problems caused by the gap in time	<input type="checkbox"/> Explains why the time gap limits the reliability of the source as evidence of what happened at the first Thanksgiving
<b>Formulate a claim</b>	<input type="checkbox"/> Irrelevant response or no response provided	<input type="checkbox"/> Makes a claim, but it is implausible	<input type="checkbox"/> Makes a claim, but it is not explicitly related to the context of the source	<input type="checkbox"/> Makes a plausible claim about a possible motivation for the depiction of the event based on the historical context of the painter
<b>Apply historical content</b>	<input type="checkbox"/> Irrelevant or no content provided	<input type="checkbox"/> Provides inaccurate or contradictory historical information	<input type="checkbox"/> Provides generalized historical information to support claim	<input type="checkbox"/> Provides specific, accurate historical information
<b>Convey historical reasoning in writing</b>	<input type="checkbox"/> No response	<input type="checkbox"/> Grammatical errors and/or awkward wording that inhibit understanding	<input type="checkbox"/> Some grammatical errors and/or instances of awkward wording that detract from effectively conveying ideas	<input type="checkbox"/> Clearly written and easy to read; few, if any, grammatical or mechanical errors
<b>OVERALL RATING</b>	<input type="checkbox"/> <b>Not Evident</b> 0 marks	<input type="checkbox"/> <b>Basic</b> 1 – 5 marks	<input type="checkbox"/> <b>Emergent</b> 6 – 9 marks (with none @ “Not Evident”)	<input type="checkbox"/> <b>Proficient</b> 10 – 12 marks (with none @ “Basic” or “Not Evident”)

## ARCH Historical Thinking Skills Rubric - Elementary

**GENERIC  
SKILLS-BASED RUBRIC**

Criteria	Close-Reading Strategies		Strategies/Procedural Concepts		Procedural Concepts	
	Sourcing	Critical Reading (Author's Craft)	Corroboration	Contextualization	Claim	Evidence
<b>4</b>	<ul style="list-style-type: none"> <li>▪ Identifies all author(s) and the original dates of a variety of primary and secondary sources.</li> <li>▪ Evaluates the reliability of the source(s), based on when and why they were written and the author's perspective.</li> </ul>	<ul style="list-style-type: none"> <li>▪ Identifies the author's viewpoint and claims, based on what is written and what the author leaves out.</li> <li>▪ Cites examples of how the author uses persuasive language, and specific words and phrases, to influence the reader.</li> </ul>	Analyzes multiple accounts of the same event or topic, noting important similarities and differences.	<ul style="list-style-type: none"> <li>▪ Applies prior and new knowledge to determine the historical setting of the source.</li> <li>▪ Uses that setting to attempt to interpret the source within that historical context, as opposed to a "present-day mindset."</li> </ul>	Formulates plausible interpretation, argument, or claim based on an evaluation of the evidence found in a variety of primary and secondary sources.	Justifies claims using appropriate, direct evidence from a variety of reliable sources.
<b>3</b>	<ul style="list-style-type: none"> <li>▪ Identifies most author(s) and the original dates of a variety of primary and secondary sources.</li> <li>▪ Examines the reliability of the source(s), based on when and why they were written and the author's perspective.</li> </ul>	<ul style="list-style-type: none"> <li>▪ Identifies the author's viewpoint and claims, based on what is written.</li> <li>▪ Identifies at least one way the author attempts to influence the reader (persuasive language, specific words and phrases, etc.).</li> </ul>	Identifies similarities and differences by comparing information and perspectives in multiple documents.	<ul style="list-style-type: none"> <li>▪ Applies prior and new knowledge to determine the historical setting of the source.</li> <li>▪ May attempt to interpret some with a "present-day mindset."</li> </ul>	Generates a reasonable interpretation, argument, or claim based on an evaluation of the evidence found in selected primary and secondary sources.	Justifies claims using some appropriate, direct evidence from a variety of reliable sources.
<b>2</b>	<ul style="list-style-type: none"> <li>▪ Identifies some author(s) and some original dates of primary and secondary sources.</li> <li>▪ Attempts to evaluate the reliability of the source(s).</li> </ul>	<ul style="list-style-type: none"> <li>▪ Attempts to identify the author's viewpoint and claim.</li> <li>▪ Attempts to identify how the author tries to influence the reader.</li> </ul>	Identifies similarities and differences in two or more sources.	Attempts to determine the historical setting of the source.	States an interpretation, argument, or claim that may or may not be based on the evidence found in selected primary and secondary sources.	Justifies claims using generalizations or limited, appropriate, direct evidence.
<b>1</b>	<ul style="list-style-type: none"> <li>▪ Identifies few author(s) and few original dates of primary and secondary sources.</li> <li>▪ Does not attempt to evaluate the reliability of the source(s).</li> </ul>	Demonstrates little to no attempt to identify the author's viewpoint or claim.	Demonstrates little to no attempt to examine documents for corroborating or conflicting evidence.	Demonstrates no attempt to understand the historical setting of the source.	Does not state an original claim, argument, or interpretation.	Does not justify or support claims using appropriate, direct evidence.

©UMBC Center for History Education, 2013. Adapted by Assessment Resource Center for History (ARCH) from the work of the Stanford History Education Group® and Bruce VanSledright, *Assessing Historical Thinking and Understanding: Innovative Ideas for New Standards*, (New York: Routledge, 2014).

## Social Studies Skills

VUS.1	The student will demonstrate skills for historical thinking, geographical analysis, economic decision making, and responsible citizenship by
a)	synthesizing evidence from artifacts and primary and secondary sources to obtain information about events in Virginia and United States history;
b)	using geographic information to determine patterns and trends in Virginia and United States history;
c)	interpreting charts, graphs, and pictures to determine characteristics of people, places, or events in Virginia and United States history;
d)	constructing arguments, using evidence from multiple sources;
e)	comparing and contrasting historical, cultural, economic, and political perspectives in Virginia and United States history;
f)	explaining how indirect cause-and-effect relationships impact people, places, and events in Virginia and United States history;
g)	analyzing multiple connections across time and place;
h)	using a decision-making model to analyze and explain the incentives for and consequences of a specific choice made;
i)	identifying the rights and responsibilities of citizenship and ethical use of material and intellectual property; and
j)	investigating and researching to develop products orally and in writing.

### Vertical Sequence of Social Studies Skill .1a

K	viewing artifacts and primary and secondary sources to develop an understanding of history;
1	using artifacts and primary and secondary sources to develop an understanding of Virginia history;
2	identifying artifacts and primary and secondary sources to understand events in American history;
3	identifying artifacts and primary and secondary sources to understand events in world cultures;
VS	analyzing and interpreting artifacts and primary and secondary sources to understand events in Virginia history;
USI	analyzing and interpreting artifacts and primary and secondary sources to understand events in United States history;
USII	analyzing and interpreting artifacts and primary and secondary sources to understand events in United States history;
C&E	analyzing and interpreting evidence from primary and secondary sources, including charts, graphs, and political cartoons;
WG	synthesizing evidence from artifacts and primary and secondary sources to obtain-information about the world's countries, cities, and environments;
WH	synthesizing evidence from artifacts and primary and secondary sources to obtain information about events in world history;
WH	synthesizing evidence from artifacts and primary and secondary sources to obtain information about events and life in world history;
VUS	synthesizing evidence from artifacts and primary and secondary sources to obtain information about events in Virginia and United States history;
GOV	planning inquiries by synthesizing information from diverse primary and secondary sources;

## STANDARD VUS.1a

The student will demonstrate skills for historical thinking, geographical analysis, economic decision making, and responsible citizenship by

a) synthesizing evidence from artifacts and primary and secondary sources to obtain information about events in Virginia and United States history;

Essential Understandings	Experiences may include but are not limited to the following:
<p>Synthesis involves combining processed information with other knowledge to logically reach a new interpretation and understanding of content.</p> <p>Primary and secondary sources enable us to examine evidence closely and to place it in a broader context.</p> <p>An artifact is an object or tool that tells us about the people from the past.</p> <p>A primary source is an artifact, document, image, or other source of information that was created during the time under study.</p> <p>A secondary source is a document, image, or other source of information that relates or discusses information originally presented elsewhere.</p> <p>Historical information may be acquired from a variety of sources:</p> <ul style="list-style-type: none"><li>• Diaries</li><li>• Interviews</li><li>• Letters</li><li>• Raw data</li><li>• Court records and transcripts</li><li>• Photographs</li><li>• Journal articles that report the findings of original research and are written by the researchers themselves</li><li>• Autobiographies</li><li>• Speeches</li><li>• Creative works (novels, plays, poems, music, art)</li><li>• Magazine and journal articles</li><li>• Nonfiction books</li></ul>	<ul style="list-style-type: none"><li>• Use a variety of information sources to do the following:<ul style="list-style-type: none"><li>○ Examine and analyze information about the social/political organization, religion, and economics of a region. Use the information gathered to gain a new and deeper understanding of the inhabitants, resources, land and water usage, transportation methods, and communications of the region.</li><li>○ Examine and analyze geographic information and demographic data. Use the information gathered to gain a new and deeper understanding of conflicts within or among regions.</li></ul></li><li>• Prepare a collection of primary and secondary sources that best relate to a unit of study. Examine the sources to do the following:<ul style="list-style-type: none"><li>○ Make and record observations about the sources and generate questions about each item.</li><li>○ Determine commonalities and patterns in the themes of the sources, as well as how the sources connect to the overarching topic of the lesson.</li><li>○ Create a graphic organizer that describes how each source depicts a specific point of view about a period of study.</li></ul></li></ul>

## OSPI-Developed Assessments for Social Studies

Office of Superintendent of Public Instruction. Social Studies. Assessments. Available at <http://www.k12.wa.us/SocialStudies/WhatAreCBAs.aspx>.

<b>ELEMENTARY</b>	<b>Civics</b>			
	You Decide*		Whose Rules*	
	<b>Economics</b>			
	Meeting Needs and Wants		People on the Move	
	<b>Geography</b>			
	Humans and the Environment	Cultural Contributions		People on the Move
	<b>History</b>			
	What's the Big Idea	Cultural Contributions	Causes of Conflict	Dig Deep
<b>MIDDLE</b>	<b>Civics</b>			
	Checks and Balances*		Constitutional Issues	
	<b>Economics</b>			
	Meeting Needs and Wants		International Relations	
	<b>Geography</b>			
	Humans and the Environment		People on the Move	
	<b>History</b>			
Enduring Cultures	Why History?	Causes of Conflict	Dig Deep	
<b>HIGH</b>	<b>Civics</b>			
	Checks and Balances*		Constitutional Issues*	
	<b>Economics</b>			
	You and the Economy	Government Revenue and Responsibility		Causes of Conflict
	<b>Geography</b>			
	Humans and the Environment		Cultural Interactions	
	<b>History</b>			
	Technology Through the Ages	Cultural Interactions	Causes of Conflict	Dig Deep
	<b>International Perspectives</b>			
	US Foreign Policy			

\*RCW 28A.230.095 specifies that school districts shall require students in the fourth or fifth grade, the seventh or eighth grade, and the eleventh or twelfth grade to complete at least one classroom-based assessment in **civics**.

**This Page Intentionally Left Blank for Your Use**

# Guidelines for Using the Simplified PBA Design Template

## I. Intended Learning Outcomes (ILOs)

- A. Be sure to identify specific ILOs to which you are aligning.
- B. When identifying what students are intended to “DO,” be sure to include subject-specific skills as appropriate (e.g., analyzing primary and secondary sources, interpreting maps, graphs, and charts, making observations, predictions, inferences, and recognizing sequences and chronologies).
- C. When identifying what students are intended to “DO,” identify the level of cognitive demand/rigor using Bloom’s revised taxonomy of cognitive behaviors or another cognitive taxonomy, such as Webb’s Depth of Knowledge (DoK). (This may be in the form of a “table of specifications.”)
- D. ILOs may be integrated (i.e., including ILOs from more than one subject).
- E. ILOs may also align to one or more transfer goals (e.g., critical thinking, responsible citizenship).

## II. Task Description--Use the RAFTS acronym to outline an authentic, appropriate performance task for students to undertake. RAFTS is a tool for teachers to use to think about essential features of a PBA. It need not be followed prescriptively, but can be a very useful starter. Here’s a brief descriptor of each element:

- A. Role--Roles might be a job or career that an adult might have, a situation in which this task might occur in an adult’s life outside of a job situation, a situation that a student might experience outside of the school, or even the most basic “role” of the student simply being a student of the discipline.
- B. Audience--The audience should be relevant to the role. For instance, if the role is as a lawyer than the “audience” in the task might be a judge or jury.
- C. Format--Format is the means by which students are to represent their thinking. Formats can be practically anything: Essay, mock legal brief, lab report, scripted dialogue, spoken dialogue recorded on an iPhone, diorama, narrated Power Point presentation, etc. Typically, the “format” is a product (creating something) and/or a performance (doing something).
- D. Task--Task is the problem they are attempting to solve or the need they are attempting to meet. The task may include specific steps that the student should undertake to complete the task.
- E. Scenario--Scenario is the authentic situation that provides context or relevance to the task. The PBAs engage students in relevant tasks that are representative of the real world and/or relevant to the discipline. Authenticity may be relevant to

jobs/careers, future responsibilities as adults, current interests of students as young people, or academic tasks relevant to the discipline.

- F. The task description is the foundation for developing a student prompt and accompanying materials for the student. **(Include the prompt and any accompanying student-facing materials for the PBA.)**

- III. **Success Criteria**--The success criteria outline expectations of student performance. In identifying success criteria, it is important to keep in mind not only the explicit knowledge and skills that students are to demonstrate, but also the relative importance of each of these. Success criteria are the foundation for developing a rubric for the PBA. (Note: A “rubric” may take the form of an analytic rubric, holistic rubric, rating scale, or checklist. Also, a rubric may be written in teacher-friendly language or in student-friendly language, depending on whether the purpose of scoring is summative or formative.) **(Include the rubric with the PBA.)**
- IV. **Instruction**--Briefly describe the instructional experiences that students will need to have experienced in order to have had the opportunity to learn the knowledge, understandings, and skills to undertake the assessment task. If this is a performance-based assessment (PBA) that is part of a project-based learning (PBL) unit, then you may simply reference that unit, if it is written up elsewhere.
- V. **Administration of the PBA**--Briefly note the required materials (including links to sites) needed for administering the assessment. Note the time needed, technology required, and if outside partnerships are utilized.
- VI. **Differentiation**--Like any other type of assessment, PBAs should allow for the participation of all students. A PBA may allow students to access the assessment through multiple entry points, while providing diverse ways of responding to the task to support accessibility. Directions for teachers for a given PBA can define appropriate, allowable supports or alternatives to facilitate accessibility while maintaining the validity of the PBA.

Subject(s) \_\_\_\_\_ Grade(s) \_\_\_\_\_

Designer(s) \_\_\_\_\_ School \_\_\_\_\_

---

I. *What intended learning outcomes are assessed through this performance-based assessment (PBA)?*

Know & Understand

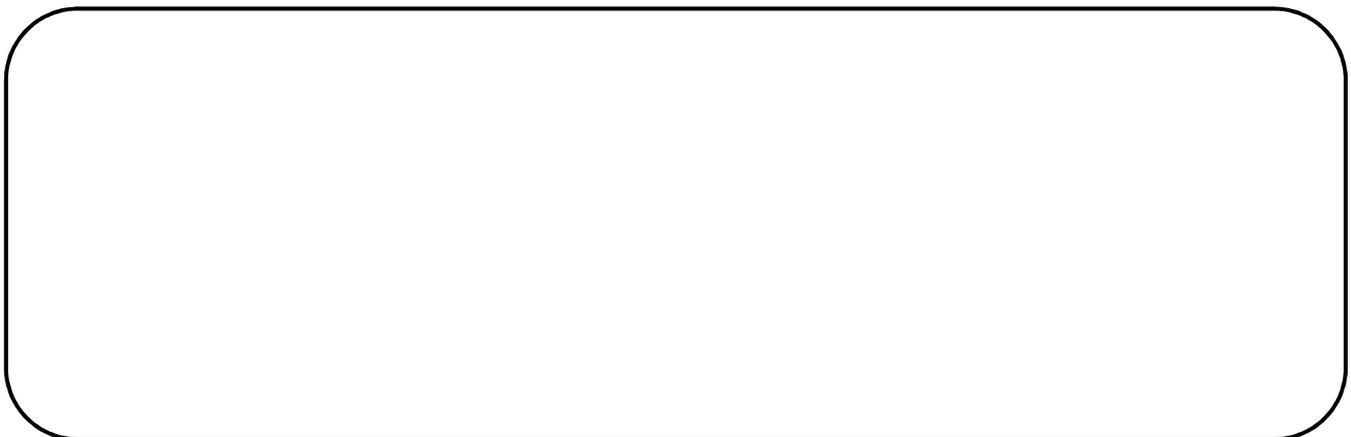
Do



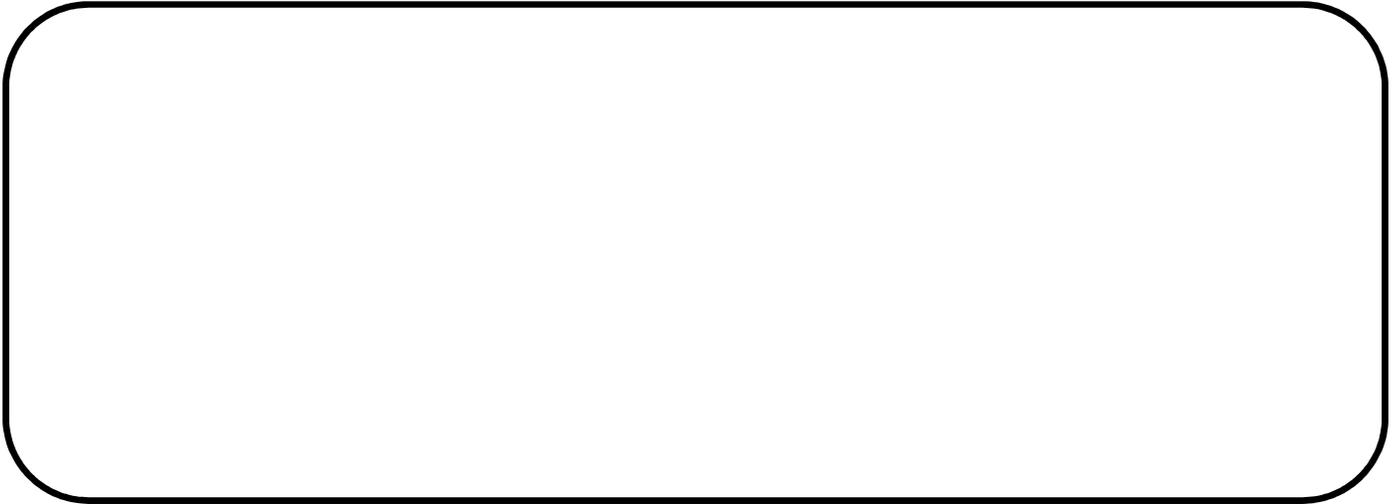
II. *Task Description (RAFTS--Role, Audience, Format, Task, Scenario)*



III. *By what criteria will students' performance and/or product be evaluated?*



IV. *What instructional experiences are necessary for students to have experienced prior to this assessment?*



V. *Directions for teachers administering this PBA (namely, time, materials, and space):*



VI. *Considerations for differentiating this PBA:*

