

SAMPLE ASSESSMENT REVIEW

A Companion to the: Assessment Review Tool

A PART OF THE ASSESSMENT TOOLKIT



RIDE Rhode Island
Department
of Education

Table of Contents

Sample Assessment for Review	Pages 3-5
Sample Annotated Alignment: Assessment Review Tool	Pages 6-7
Science Depth-of-Knowledge	Pages 8-10
Sample Assessment Scoring Criteria and Student Work for Review	Pages 11-18
Sample Annotated Rubric and Scoring Guide and Criteria: Assessment Review Tool	Page 19
Sample Assessment Administration Procedures for Review	Pages 20-21
Sample Annotated Fair and Unbiased: Assessment Review Tool	Pages 22-23
Understanding Accommodations	Page 24
Recommendations	Page 25

Sample Assessment for Review

The Blizzard of 1993

Task with Student Directions

Grade 5-8 Performance Task

© 1997-2005

Contributed to SRI International Performance Assessment Links in Science

by: New York State Education Department (NYSED)

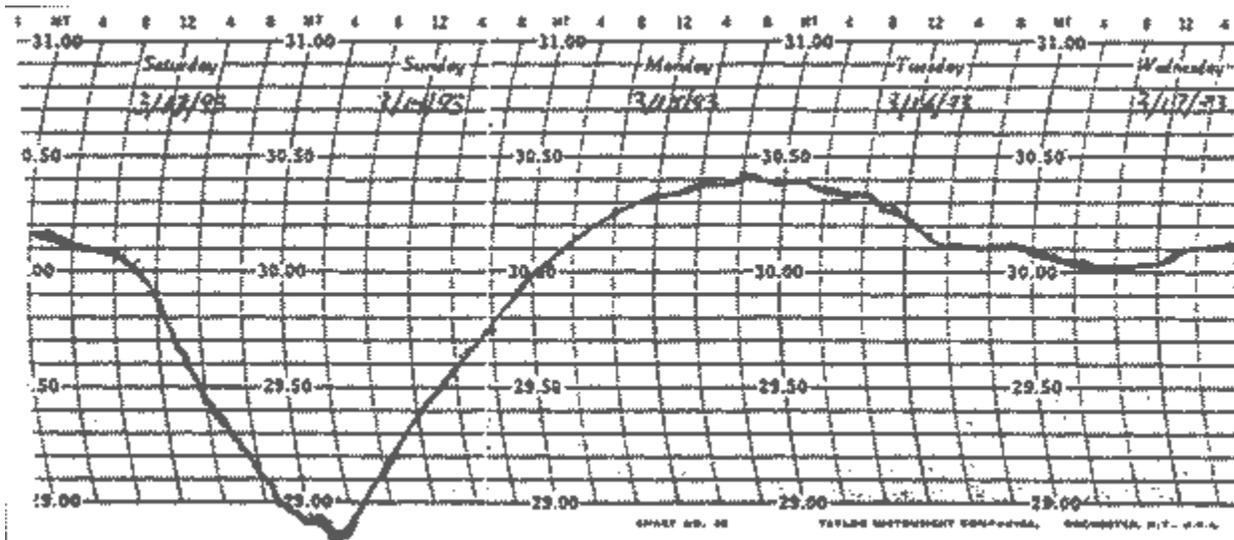
NYS Alternative Assessment in Science Project (1996)

Task:

In this activity, you will be reading and interpreting the information from a barograph from a major winter storm.

Background:

Over the weekend of March 13 & 14, 1993, over two (2) feet of snow fell over much of New York State. A barograph continuously recorded the air pressure during the storm. The chart shows the air pressures as recorded by that barograph.



1. In the space below, make a table of the barometric pressures from the barograph at four hour intervals. Start the table at 8:00 am on Saturday and end with 8:00 am on Sunday. Record the barometric pressures to the nearest tenth of an inch.

2. A very distinct cold front passed through this region. Based on the graph, on which day and time did it likely arrive?

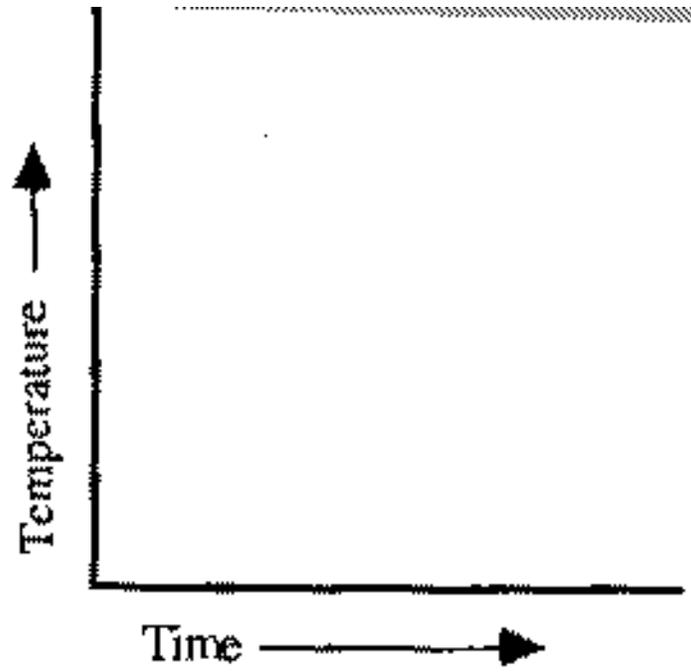
3. What observations from the graph led you to your conclusion in question #2?

4. Based on what you know about the relationship between wind velocity and air pressure, during which four hour period was the wind velocity likely the greatest?

5. Using complete sentences, describe how you used the barograph to determine your answer to question #4.

6. Using complete sentences, describe the relationship between air pressure and air temperature between 8:00 PM on Saturday and 8:00 am on Sunday.

7. Draw a line graph of the temperature pattern during the time period described in question #6.



Sample Annotated Alignment: Assessment Review Tool

A high quality teacher-created assessment should be ... Aligned			
Part 2: Alignment			
Identify the SLO for which this assessment is being considered: 6 th Grade Earth and Space Science			
Identify the standards included in the SLO:			
SLO Standards	Evaluated by Assessment		
	Fully	Partially	No
ESS1 (5-6)- 2c developing models to explain how humidity, temperature, and altitude affect air pressure and how this affects local weather.	X		
INQ Construct 8: Use accepted methods for organizing, representing, and manipulating data (DOK2)	X		
INQ Construct 12: Use evidence to support and justify interpretations and conclusions or explain how the evidence refutes the hypothesis (DOK 3)		X	
Scientific Inquiry GSE: ESS1 (5-6) 4a: explaining how differential heating and convection affect Earth's weather patterns			X
4b: describing how differential heating of the oceans affects ocean currents which in turn influence weather and climate			X
4c: explaining the relationship between differential heating/convection and the production of winds			X
Identify additional standards measured by this assessment that are not included in the SLO:			
Additional Standards	Check if scores can be partitioned		
1.			
2.			
3.			
4.			
5.			
Describe the content knowledge/concepts assessed: Examine weather data, draw conclusions based on the evidence in the barograph, and make predictions based on it.			
Describe the skills/performance assessed: Create a graph, read data from a graph			
Are there a sufficient number of items or tasks on the assessment for each standard being assessed? If no, please identify which standards are underrepresented.			
The assessment includes two process questions, expecting students to interpret information and create a table and a graph. In addition, there are 5 short answer questions that expect students to interpret the graph provided and to provide evidence from the graph. However, the assessment does not have <u>any</u> questions/tasks that are specific to the ESS GSEs. Therefore, there are not a sufficient number of items on this assessment to target the standard being assessed.			

Scientific Inquiry includes more than what is identified for the assessment and incorporate higher levels of DOK (e.g., designing, conducting & critiquing an investigation is a DOK 4). Therefore, the DOK range for Scientific Inquiry is DOK 1-4.

In addition, although this assessment asks students to draw a conclusion and make predictions, there is only one correct response and students are only asked for a sentence to explain their thinking.

An implication of this comparison is that another assessment may be needed to assess the upper range of Depth-of-Knowledge expectations in the standards.

<p>To what extent do you see a strong content match between the task or items and the corresponding state-adopted standards?</p> <p><input type="checkbox"/> Full match – all tasks or items fully address or exceed the relevant skills and knowledge described in the corresponding state-adopted standard(s)</p> <p><input type="checkbox"/> Close/Partial match – most tasks or items address or partially address the relevant skills and knowledge described in the corresponding state-adopted standard(s)</p> <p><input checked="" type="checkbox"/> Minimal match – some tasks or items match some relevant skills and knowledge described in the corresponding state-adopted standard(s)</p> <p><input type="checkbox"/> No match – tasks or most items are not related to the skills and knowledge described in the corresponding state-adopted standard(s)</p>	
<p>Are the items or tasks reviewed as cognitively complex as the state-adopted standards? Use the definitions below to select your rating.</p> <p><input type="checkbox"/> More complex – most items or the tasks reviewed are at a higher level of complexity than the range indicated for the state-adopted standard(s)</p> <p><input type="checkbox"/> Similarly complex – most items or the task reviewed are at a similar level of complexity as the state-adopted standard(s)</p> <p><input checked="" type="checkbox"/> Less complex – most items or the task reviewed are at a lower level of complexity than the state-adopted standard(s)</p>	
<p>Comments/Suggestions for Improvement</p>	
<p>Provide evidence to support your responses:</p> <p>This task is a DOK 2-3 Range. DOK Level 4 requires complex reasoning, experimental design and planning and probably will require an extended period of time, either for the science investigation required by and objective, or for carrying out the multiple steps of an assessment item. This task does not fulfill those criteria.</p> <p>The assessment aligns to some of the scientific inquiry expectations and allows for some processing of information. Expanding this assessment to include all of the related GSEs (ESS1 (5-6) 4), and/or research about the content expected in the GSEs, and to make connections between the data viewed and the content would strengthen this assessment.</p> <p>Finally, the students are not given clear data about air temperature, making it difficult to identify and chart the relationship between air pressure readings and air temperature.</p>	

Depth-of-Knowledge

There are many frameworks for measuring cognitive demand. This document refers to Webb's Depth of Knowledge Framework (2002), which outlines four levels of cognitive demand that are applicable to all content levels:

Level 1 is Recall and is characterized by simple retelling or recitation of facts or a procedure.

Level 2 is Skill/Concept and necessitates some type of decision-making. The response to a prompt will not be automatic and will require more than one step for the student to arrive at the answer.

Level 3 is Strategic Thinking. This is where reasoning becomes more complex and demanding. Tasks of this variety require greater planning, abstraction, evidence, and justification from the student. A student engaged in Level 3 is often required to form a hypothesis or conjecture.

Level 4 is Extended Thinking and manifests itself in tasks that require an extended period of time utilizing complex thinking and planning. Level 4 tasks compel students to make connections within a discipline and/or to other disciplines. More than likely, there are multiple solutions to a problem and multiple pathways for attaining a solution. Level 4 tasks are not typically found in large-scale assessments as they usually require multiple days of thought and consideration by the student. Students should be applying what they know to new situations to come up with complex answers and justifications.

It is important to note that Depth of Knowledge levels are not discrete but rather they are on a continuum.

Descriptors of DOK Levels for Science

(based on Webb, March 2002 and TIMSS Science Assessment Framework, 2003)

Level 1 Recall and Reproduction requires recall of information, such as a fact, definition, term, or a simple procedure, as well as performing a **simple** science process or procedure. Level 1 only requires students to demonstrate a rote response, use a well-known formula, follow a set procedure (like a recipe), or perform a clearly defined series of steps. A “simple” procedure is well-defined and typically involves only **one-step**. Verbs such as “identify,” “recall,” “recognize,” “use,” “calculate,” and “measure” generally represent cognitive work at the recall and reproduction level. Simple word problems that can be directly translated into and solved by a formula are considered Level 1. Verbs such as “describe” and “explain” could be classified at different DOK levels, depending on the complexity of what is to be described and explained. A student answering a Level 1 item either knows the answer or does not: that is, the answer does not need to be “figured out” or “solved.” In other words, if the knowledge necessary to answer an item automatically provides the answer to the item, then the item is at Level 1. If the knowledge necessary to answer the item does not automatically provide the answer, the item is at least at Level 2.

Level 2 Skills and Concepts includes the engagement of some mental processing beyond recalling or reproducing a response. The content knowledge or process involved is **more complex** than in level 1. Items require students to make some decisions as to how to approach the question or problem. Keywords that generally distinguish a Level 2 item include “classify,” “organize,” “estimate,” “make observations,” “collect and display data,” and “compare data.” These actions imply **more than one step**. For example, to compare data requires first identifying characteristics of the objects or phenomenon and then grouping or ordering the objects. Level 2 activities include making observations and collecting data; classifying, organizing, and comparing data; and organizing and displaying data in tables, graphs, and charts. **Some action verbs, such as “explain,” “describe,” or “interpret,” could be classified at different DOK levels, depending on the complexity of the action.** For example, interpreting information from a simple graph, requiring reading information from the graph, is a Level 2. An item that requires interpretation from a complex graph, such as making decisions regarding features of the graph that need to be considered and how information from the graph can be aggregated, is at Level 3.

Level 3 Strategic Thinking requires reasoning, planning, using evidence, and a more **complex and abstract** cognitive demands than Levels 1 and 2. The complexity does not result only from the fact that there could be multiple answers, but because the multi-step task requires **more demanding reasoning**. In most instances, requiring a simple explanation of a word or two is a Level 2, whereas tasks requiring students to explain their thinking are at Level 3; An activity that has more than one possible answer and requires students to justify the response they give would most likely be a Level 3. Experimental designs in Level 3 typically involve more than one dependent variable. Other Level 3 activities include drawing conclusions from observations; citing evidence and developing a logical argument; explaining phenomena in terms of concepts; and using concepts to solve non-routine problems.

Level 4 Extended Thinking requires **high cognitive demand** and is **very complex**. Students are required to make several connections—relate ideas *within* the content area or *among* content areas—and select or devise one approach among many alternatives on how the situation can be solved. Many on-demand assessment instruments will not include any assessment activities that could be classified as Level 4. However, standards, goals, and objectives can be stated in such a way as to expect students to perform extended thinking. “Develop generalizations of the results obtained and the strategies used and apply them to new problem situations,” is an example of a Grade 8 objective that is a Level 4. Many, but not all, performance assessments and open-ended assessment activities requiring significant thought will be at a Level 4. Level 4 requires complex reasoning, experimental design and planning, and **probably will require an extended period of time** either for the science investigation required by an objective, or for carrying out the multiple steps of an assessment item. However, the extended time period is not a distinguishing factor if the required work is only repetitive and does not require applying conceptual understanding and higher-order thinking. For example, if a student has to take the water temperature from a river each day for a month and then construct a graph, this would be classified as a Level 2 activity. However, if the student conducts a river study that requires taking into consideration a number of variables, this would be a Level 4.

Detailed Descriptions of Depth of Knowledge Levels for Science

K. Hess, Center for Assessment, updated 2005)

Level 1 Recall of Information	Level 2 Skills/Concepts	Level 3 Strategic Thinking	Level 4 Extended Thinking
Examples represent, but do not constitute all Level 1 science performances:	Examples represent, but do not constitute all Level 2 science performances:	Examples represent, but do not constitute all Level 3 science performances:	Examples represent, but do not constitute all Level 4 science performances:
<p>Recall or recognize a fact, term, definition, simple procedure (such as one step), or property</p> <p>Demonstrate a rote response</p> <p>Use a well-known formula</p> <p>Represent in words or diagrams a scientific concept or relationship</p> <p>Provide or recognize a standard scientific representation for simple phenomenon</p> <p>Perform a routine procedure, such as measuring length</p> <p>Perform a simple science process or a set procedure (like a recipe)</p> <p>Perform a clearly defined set of steps</p> <p>Identify, calculate, or measure</p>	<p>Specify and explain the relationship between facts, terms, properties, or variables</p> <p>Describe and explain examples and non-examples of science concepts</p> <p>Select a procedure according to specified criteria and perform it</p> <p>Formulate a routine problem given data and conditions</p> <p>Organize, represent, and compare data</p> <p>Make a decision as to how to approach the problem</p> <p>Classify, organize, or estimate</p> <p>Compare data</p> <p>Make observations</p> <p>Interpret information from a simple graph</p> <p>Collect and display data</p> <p>NOTE: If the knowledge necessary to answer an item <u>does not</u> automatically provide the answer, then the item is at least a Level 2. Most actions imply more than one step.</p>	<p>Interpret information from a complex graph (such as determining features of the graph or aggregating data in the graph)</p> <p>Use reasoning, planning, and evidence</p> <p>Explain thinking (beyond a simple explanation or using only a word or two to respond)</p> <p>Justify a response</p> <p>Identify research questions and design investigations for a scientific problem</p> <p>Use concepts to solve non-routine problems/more than one possible answer</p> <p>Develop a scientific model for a complex situation</p> <p>Form conclusions from experimental or observational data</p> <p>Complete a multi-step problem that involves planning and reasoning</p> <p>Provide an explanation of a principle</p> <p>Justify a response when more than one answer is possible</p> <p>Cite evidence and develop a logical argument for concepts</p> <p>Conduct a designed investigation</p> <p>Research and explain a scientific concept</p> <p>Explain phenomena in terms of concepts</p>	<p>Select or devise approach among many alternatives to solve problem</p> <p>Based on provided data from a complex experiment that is novel to the student, deduct the fundamental relationship between several controlled variables.</p> <p>Conduct an investigation, from specifying a problem to designing and carrying out an experiment, to analyzing its data and forming conclusions</p> <p>Relate ideas <i>within</i> the content area or <i>among</i> content areas</p> <p>Develop generalizations of the results obtained and the strategies used and apply them to new problem situations</p> <p>NOTE: Level 4 activities often require an extended period of time for carrying out multiple steps; however, time alone is not a distinguishing factor if skills and concepts are simply repetitive over time.</p>
<p>NOTE: If the knowledge necessary to answer an item automatically provides the answer, it is a Level 1.</p>			

Sample Assessment Scoring Criteria and Student Work for Review

The Blizzard of 1993																	
Rubric																	
Grade 5-8 Performance Task Contributed by: New York State Education Department (NYSED)NYS Alternative Assessment in Science Project (1996) Scoring Rubric - The Blizzard of 1993"																	
Maximum Score - 11 points																	
1. Data Table	4 points total																
<p><u>Standard:</u> The student will construct a data table from the information on the barograph.</p> <p><u>Criteria:</u></p> <ul style="list-style-type: none"> ▪ Allow 1 point for setting up an accurate and usable data table. ▪ Allow 2 points for 5, 6, or 7 correct values entered in the table from the barograph information. ▪ Allow 1 point for 3 or 4 correct values entered in the table from the barograph information. <p style="text-align: center;">Table 1: Air Pressure Readings in Inches of Mercury Between March 13 & 14,1993 for NY State</p> <table border="1" style="margin: 10px auto; border-collapse: collapse; text-align: center;"> <thead> <tr> <th style="padding: 5px;">Time</th> <th style="padding: 5px;">8 am</th> <th style="padding: 5px;">12 PM</th> <th style="padding: 5px;">4 PM</th> <th style="padding: 5px;">8 PM</th> <th style="padding: 5px;">MT</th> <th style="padding: 5px;">4 am</th> <th style="padding: 5px;">8 am</th> </tr> </thead> <tbody> <tr> <td style="padding: 5px;">Pressure (inHg)</td> <td style="padding: 5px;">30.1</td> <td style="padding: 5px;">29.9</td> <td style="padding: 5px;">29.4</td> <td style="padding: 5px;">29.1</td> <td style="padding: 5px;">28.9</td> <td style="padding: 5px;">28.9</td> <td style="padding: 5px;">29.1</td> </tr> </tbody> </table> <ul style="list-style-type: none"> ▪ Allow 0 points for poorly constructed table or less than two correct values entered in the data table. ▪ Allow 1 point for ALL values entered in the table are given to the nearest tenth. 		Time	8 am	12 PM	4 PM	8 PM	MT	4 am	8 am	Pressure (inHg)	30.1	29.9	29.4	29.1	28.9	28.9	29.1
Time	8 am	12 PM	4 PM	8 PM	MT	4 am	8 am										
Pressure (inHg)	30.1	29.9	29.4	29.1	28.9	28.9	29.1										
2. Cold Front Data	1 point total																
<p><u>Standard:</u> The student will accurately pinpoint the date and time of the major cold front.</p> <p><u>Criteria:</u> Allow 1 point for indicating that the cold front arrived between 12:00 am and 4:00 am on Sunday, March 14.</p>																	
3. Evidence of Passing Front	1 point total																
<p><u>Standard:</u> The student will cite evidence from the barograph that explains the passing of the cold front.</p> <p><u>Criteria:</u></p> <ul style="list-style-type: none"> ▪ Allow 1 point for indicating a sharp change in direction of the barograph line or the "spike" in the pressure line. 																	

4. Wind Velocity Data**1 point total**

Standard: The student will accurately pinpoint the data and time of the greatest wind velocity during the storm.

Criteria:

- Allow 1 point for indicating that the period of the greatest wind velocity was Sunday 2:00 am to 8:00 am.
-

5. Evidence of Wind Velocity**2 points total**

Standard: The student will cite evidence from the barograph that indicates high wind velocity.

Criteria:

- Allow 2 points for indicating that the period of greatest wind velocity will be the area on the graph that has the steepest slope, using a complete sentence.
 - Allow 1 point for a correct answer, but not in a complete sentence.
 - Allow 0 points for an incorrect answer even if it is in a complete sentence.
-

6. Temperature and Air Pressure**2 points total**

Standard: The student will explain the relationship between air pressure and air temperature.

Criteria:

Allow 2 points for indicating:

- Air pressure and temperature show an inverse relationship in a complete sentence.
- Allow 1 point for a correct answer, but not in a complete sentence.
- Allow 0 points for an incorrect answer even if it is in a complete sentence.

Allow 1 point for indicating one of the variables correctly.

7. Graph**1 point total**

Standard: The student will sketch a graph showing the change in temperature during a blizzard.

Criteria:

- Allow 1 point for a line drawn on the graph that shows the temperature rising or steady and then falling sharply.
-

Total Score _____

Total possible score - 12 points

STUDENT #1

1. Data Table			
Usable and accurate table	0	①	
Correct values entered	0	1	②
Values to the nearest tenth	0	①	
2. Cold Front Data	0	①	
3. Evidence of Passing Front	0	①	
4. Wind Velocity Data	0	①	
5. Evidence of Wind Velocity	0	1	②
6. Temperature and Air Pressure	0	1	②
7. Graph	0	①	
Total Score			<u>12</u>
	(Total possible score - 12 points)		

1. In the space below, make a table of the barometric pressures from the barograph at four hour intervals. Start the table at 8:00 am on Saturday and end with 8:00 am on Sunday. Record the barometric pressures to the nearest tenth of an inch.

Air Pressure on March 13+14 1993

Time	8 am	12 pm	4 pm	8 pm	MT	4 am	8 am
Air Pressure (in)	30.1	29.9	29.5	29.1	28.9	28.9	29.1

2. A very distinct front passed through this region. On which day and time did it arrive?

2:00 am on Sunday

3. What observations from the graph led you to your conclusion in question #2?

at about 2:00 am on Sunday the barograph recorded very rapid change in air pressure from going down to going up.

4. During which four hour period was the wind velocity the greatest?

...
about 4:00 to 8:00 am on Sunday

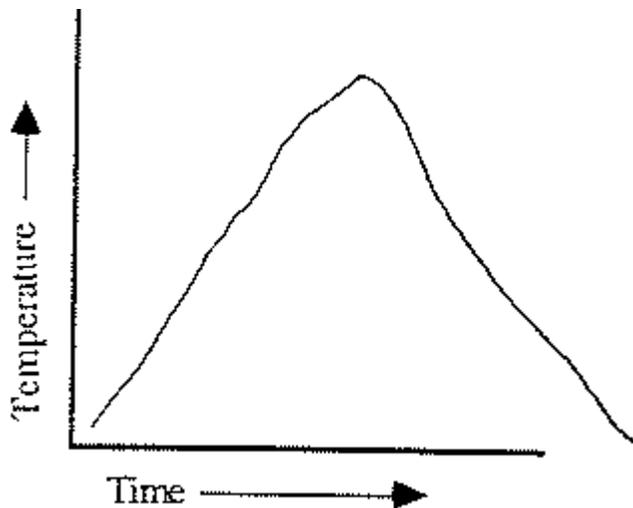
5. Using complete sentences, describe how you used the barograph to determine your answer to question #4.

The wind would have been the strongest at this time because this is when the barograph went up the fastest.

6. Using complete sentences, describe the relationship between air pressure and air temperature between 8:00 pm on Saturday and 8:00 am on Sunday.

As the air pressure went down the air temperature would most likely have gone up and the reverse.

Draw a line graph of the likely temperature pattern during the time period described in question #6.



STUDENT #2

1. Data Table	0	(1)	
Usable and accurate table	0	(1)	2
Correct values entered	0	(1)	
Values to the nearest tenth	0	(1)	
2. Cold Front Data	0	(1)	
3. Evidence of Passing Front	0	(1)	
4. Wind Velocity Data	(0)	1	
5. Evidence of Wind Velocity	(0)	1	2
6. Temperature and Air Pressure	0	1	(2)
7. Graph	0	(1)	
Total Score		8	
	(Total possible score - 12 points)		

1. In the space below, make a table of the barometric pressures from the barograph at four hour intervals.

Start the table at 8:00 am on Saturday and end with 8:00 am on Sunday. Record the barometric pressures to the nearest tenth of an inch.

Time	Air Pressure
8:00 AM	30.1 inches.
12:00 PM	29.9 "
4:00 PM	29.5 "
8:00 PM	29.2 "
Midnight.	28.8 "
4:00 AM	28.9 "
8:00 AM	29.3 "

2. A very distinct front passed through this region. On which day and time did it arrive?

3:00 AM on Sunday

3. What observations from the graph led you to your conclusion in question #2?

This is when the air pressure is lowest. Then it went up.

4. During which four hour period was the wind velocity the greatest?

Monday from 4 PM to 8 PM.

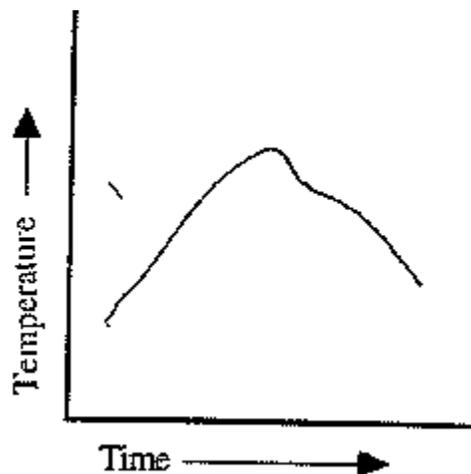
5. Using complete sentences, describe how you used the barograph to determine your answer to question #4.

This was the time when the air pressure was the highest.

6. Using complete sentences, describe the relationship between air pressure and air temperature between 8:00 pm on Saturday and 8:00 am on Sunday.

On Saturday the air pressure was going down because the air was warm. On Sunday air pressure went back up and it got cold.

Draw a line graph of the likely temperature pattern during the time period described in question #6.



STUDENT #3

1. Data Table	0	(1)	
Usable and accurate table			
Correct values entered	(0)	1	2
Values to the nearest tenth	(0)	1	
2. Cold Front Data	0	(1)	
3. Evidence of Passing Front	0	(1)	
4. Wind Velocity Data	0	1	
5. Evidence of Wind Velocity	(0)	1	2
6. Temperature and Air Pressure	(0)	1	2
7. Graph	(0)	1	
Total Score		<u>3</u>	
			(Total possible score - 12 points)

1. In the space below, make a table of the barometric pressures from the barograph at four hour intervals.

Start the table at 8:00 am on Saturday and end with 8:00 am on Sunday. Record the barometric pressures to the nearest tenth of an inch.

Pressure	8:00	12:00	4:00	8:00	12:00	4:00	8:00
	30.10	29.90	29.50	29.20	less than 29.00	less than 29.00	28.20

2. A very distinct front passed through this region. On which day and time did it arrive?

On Sunday, a cold front passed at 2:00 AM.

3. What observations from the graph led you to your conclusion in question #2?

The pressure was lowest at 2:00 AM.

4. During which four hour period was the wind velocity the greatest?

It was greatest between 12:00 and 4:00 on Sunday.

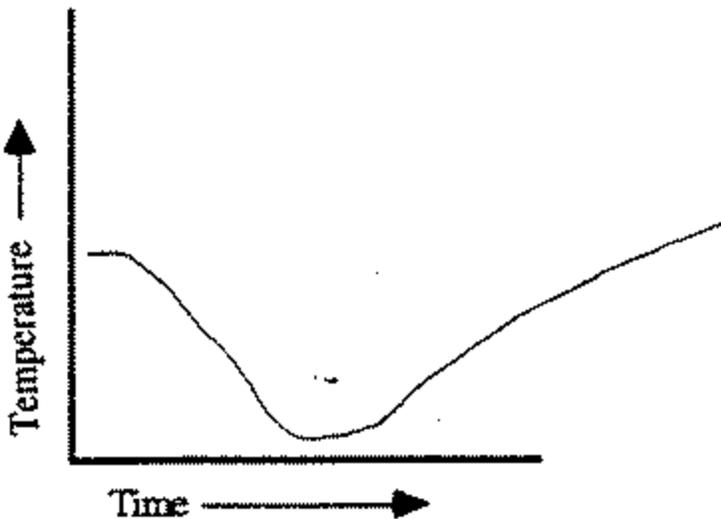
5. Using complete sentences, describe how you used the barograph to determine your answer to question #4.

Wind is highest when the pressure is low.

6. Using complete sentences, describe the relationship between air pressure and air temperature between 8:00 pm on Saturday and 8:00 am on Sunday.

The air pressure and temperature both were low.

Draw a line graph of the likely temperature pattern during the time period described in question #6.



Sample Annotated Rubric/Scoring Guidelines and Criteria: Assessment Review Tool

A high quality assessment should be ... Scored using Clear Guidelines and Criteria
Part 3: Rubric/Scoring Guide
<p>Scoring Guide to be used with the assessment:</p> <p><input type="checkbox"/> Generalized Rubric (e.g., for writing an argument, for all science labs, etc.)</p> <p><input type="checkbox"/> Task-specific Rubric (only used for the particular task)</p> <p><input checked="" type="checkbox"/> Scoring Guidelines (e.g., checklist with score points for each part)</p> <p><input checked="" type="checkbox"/> Answer Key, Scoring Template, Computerized or Machine Scored</p> <p><input type="checkbox"/> Anchor Papers (student samples at each score point)</p> <p><input type="checkbox"/> Teacher Observation Sheet/Observation Checklist</p>
<p>Are the rubric/scoring criteria aligned to the state-adopted standards? If no, please explain.</p> <p>The scoring criteria analyze students' ability to display and interpret data, citing evidence, but only hints at making predictions.</p>
<p>Are the score categories clearly defined and coherent across performance levels? If no, please explain.</p> <p>The checklist with score points is clearly defined for each question. The scoring guide appears to confuse the understanding of the content information with using appropriate English conventions. Specifically, in the Evidence of Wind Velocity, the scoring guideline states:</p> <ul style="list-style-type: none"> • Allow 2 points for indicating that the period of greatest wind velocity will be the area on the graph that has the steepest slope, using a complete sentence. • Allow 1 point for a correct answer, but not in a complete sentence. • Allow 0 points for an incorrect answer even if it is in a complete sentence.
<p>Do the rubric/scoring criteria address all of the demands within the task or item? If no, please explain.</p> <p>The scoring criteria address all of the demands expected within the task.</p>
<p>Based on your review of the rubric/scoring criteria, would the scoring rubric most likely lead different raters to arrive at the same score for a given response? If no, please explain.</p> <p>The scoring criteria are specific and will lead teachers to the same point value. However, the use of complete sentences may be problematic for science teachers scoring student work.</p>
<p>Consider how long it will likely take the teacher(s) to score each assessment. Is this practical given the number of students and the type of assessment?</p> <p>The scoring of this assessment is practical and will require a short amount of time if scored as correct or incorrect. However, the assessment provides an opportunity for teachers to identify student misconceptions and areas of need. Therefore, a teacher could spend additional time recording specific comments that will assist students with having a greater conceptual understanding.</p>
<p>Are student work samples available that illustrate different levels of performance available? If not, explain what additions are needed.</p> <p>Student work samples are included that clearly demonstrate a continuum of understanding, including a student with mastery of the content, a student with some mastery, and a student that lacks mastery.</p>
Comments/Suggestions for Improvement
<p>Provide evidence from to support your responses:</p> <p>The scoring criteria clearly address each aspect of the task and what a student is expected to understand. However, given that the assessment is not a research or essay assignment that expects appropriate written conventions, it should be modified so that it addresses the content and not English conventions.</p>

Sample Assessment Administration Procedures for Review

The Blizzard of 1993

Administration Procedures

Grade 5-8 Performance Task
Contributed by: New York State Education Department (NYSED)
NYS Alternative Assessment in Science Project (1996)

DESCRIPTION:

Students will read and interpret the information from a barograph from a major winter storm.

This task assesses students' abilities to interpret and analyze graphs, construct data tables and graphs, generalize, infer, apply knowledge of meteorology, and explain scientific relationships.

This task is designed to take students approximately 20 minutes to complete.

OVERALL TASK CONTENT AREA:

Earth and Space Science

SPECIFIC KNOWLEDGE AREAS:

Structure of the earth system

PERFORMANCE EXPECTATIONS:

- organizing and representing data
- formulating conclusions from investigational data
- applying scientific principles to develop explanations

GENERAL INSTRUCTIONS TO THE TEACHER:

This task is designed to take approximately 20 minutes to complete.

Students will be working individually during this exercise.

Students should be ready to work as soon as the period begins. The materials should be set out at each lab station, if possible. A central supply area, if needed, should be easily accessible. All supplies should be clearly labeled.

MATERIALS FOR "THE BLIZZARD OF 1993":

At each station students should have:

- barograph from March 13-19, 1993
- ruler or straight edge

ADVANCE PREPARATION:

None

SAFETY:

- Be careful.
- Teachers and students should always exercise appropriate safety precautions and utilize appropriate laboratory safety procedures and equipment when working on science performance tasks.

EXTENSIONS/MODIFICATIONS:

Students could examine other barographs from other blizzards.

Sample Annotated Fair and Unbiased: Assessment Review Tool

A high quality performance assessment should be...Fair and Unbiased	
Part 4: Fair and Unbiased (the areas below should be discussed relative to the needs of ELLs, gifted and talented students, and students with disabilities)	
To what extent are the items or tasks visually clear and uncluttered (e.g., appropriate white space and/or lines for student responses, graphics and/or illustrations are clear and support the test content, the font size seems appropriate for the students)? Formatting is:	
<input type="checkbox"/> Visually clear and uncluttered <input checked="" type="checkbox"/> Somewhat clear and uncluttered <input type="checkbox"/> Unclear, cluttered, and inappropriate for students	
Provide an explanation of your response, if needed: The questions are clearly formatted and easy to read; however, the graph that supplies the data is somewhat difficult to read and appears to have been distorted when scanned into an electronic format. In addition, the units are unmarked and, as a result, the air pressure could easily be confused with temperature. Furthermore, at this grade level, the blank graph should contain lines for a title and labels for the unit of measurement on each axis. Finally, the axis on the graph should not contain arrows, which suggests a continuous and constant relationship between temperature and time.	
Are directions for the items or tasks presented in as straightforward a manner as possible for a range of learners? If no, identify problematic items/tasks and provide suggestions for improvement. All directions and items are very direct and straightforward regarding what students are asked to demonstrate.	
Is the vocabulary and context(s) presented free from cultural or other unintended bias? If no, identify problematic items/tasks and provide suggestions for improvement. Although the basis of the data is a blizzard, in which some students may not have had direct experience with, the tasks asks students to gather and evaluate content specific vocabulary and data and therefore, the assessment can be considered to be without bias.	
Does the assessment use appropriate levels of academic language for the grade and content area? If no, identify problematic areas and provide suggestion for improvement. All academic language is basic and appropriate for the grade level and content area.	
Accommodations are commonly categorized in five ways: presentation, response, setting, timing and scheduling, and linguistics. In the table below, identify any additional accommodations that should be provided to ensure that English Learners and/or Students with Disabilities can access the content represented by the task or items.	
<input checked="" type="checkbox"/> Presentation Accommodations – Allow students to access information in ways that do not require them to visually read standard print. These alternate modes of access are auditory, multi-sensory, tactile, and visual.	This assessment can be read to students.
<input checked="" type="checkbox"/> Response Accommodations —Allow students to complete activities, assignments, and assessments in different ways or to solve or organize problems using some type of assistive device or organizer.	Written responses to this assessment can be scribed. However, the completion of graphs/charts must be completed by the student.

Identifying potential bias will help to make determinations about any concepts that may need to be pre-taught.

These accommodations were not identified on the assessment; however, these accommodations have been identified as acceptable as they do not impact the student responses.

<p><input checked="" type="checkbox"/> Setting Accommodations—Change the location in which a test or assignment is given or the conditions of the assessment setting.</p>	<p>Students can take this assessment in a setting that is most appropriate.</p>
<p><input type="checkbox"/> Timing and Scheduling Accommodations—Increase the allowable length of time to complete an assessment or assignment and perhaps change the way the time is organized.</p>	
<p><input type="checkbox"/> Linguistic Accommodations—Allow English language learners (ELLs) to access academic construct measured by reducing the linguistic load of an assessment. The accommodation is based on an ELL’s limited English language proficiency, which is different than an accommodation based on a student’s disability or a cognitive need</p>	

*Please reference “Defining Features of Academic Language in WIDA’s Standards” (See RIDE website)

<p>Comments/Suggestions for Improvement</p>
<p>Provide evidence from to support your responses: The test specifically identifies that it is limited to 20 minutes and no accommodations are written into the teacher directions. However a teacher could clearly make accommodations to assist students that would not diminish the accuracy of the information interpreted and displayed.</p>

Understanding Accommodations

Presentation Accommodations – Allow students to access information in ways that do not require them to visually read standard print. These alternate modes of access are auditory, multi-sensory, tactile, and visual.

- ▶ ***Example: text read aloud vs. text read independently***

Response Accommodations—Allow students to complete activities, assignments, and assessments in different ways or to solve or organize problems using some type of assistive device or organizer.

- ▶ ***Example: dictating response as the teacher scribes***

Setting Accommodations—Change the location in which a test or assignment is given or the conditions of the assessment setting.

- ▶ ***Example: sitting alone rather than in a group while responding to the task***

Timing and Scheduling Accommodations—Increase the allowable length of time to complete an assessment or assignment and perhaps change the way the time is organized.

- ▶ ***Example: administering the assessment in the morning when the student is more alert***

Linguistic Accommodations—Allow English language learners (ELLs) to access academic construct measured by reducing the linguistic load of an assessment. The accommodation is based on an ELL's limited English language proficiency, which is different than an accommodation based on a student's disability or a cognitive need.

- ▶ ***Example: allowing the use of a bilingual dictionary; orally translating the text***

Recommendations for this assessment:

- This assessment can be used for this SLO without revisions**
- This assessment can be used for this SLO with minor revisions** (explain below)
- This assessment can be used for this SLO with significant revisions** (explain below)
- This assessment should not be used for this SLO** (explain below)

Comments/Suggestions for Revisions

If used as a formative assessment within the classroom, it is recommended that the barograph be redone so that it is clearer for students. This assessment does not assess all of the expectations and cognitive rigor in the identified SLO. This assessment appears to be a worthwhile formative assessment that will provide information about students understanding of reading and interpreting weather data, as well as organizing and displaying it in a table or graph.