

# **Goal #2: Understand Elements of Assessment**

Formative, Interim and  
Summative Assessments

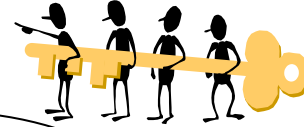
Target Types (Depth of  
Knowledge)

Target-Method- Match

Validity, reliability and Fairness

Formats

# KEYS TO QUALITY CLASSROOM ASSESSMENT



## Accurate Assessment

**Key 1: Clear Purposes**  
Why assess?  
Who will use the results?  
What will they use the results to do?

**Key 2: Clear Targets**  
What are the learning targets?  
Are they clear?  
Are they good?

**Key 3: Sound Design**  
What method?  
Quality questions?  
Sampled how?  
Avoid bias how?

## Effectively Used

**Key 4: Effective Communication**  
How manage information?  
How report? To whom?

## Key 5: Student Involvement

Students are users, too

Students can track progress and communicate, too

Students need to understand targets, too

Students can assess, too

**Key 3: Sound Design**  
**Possible Assessment Methods and Match to Targets**

**Selected Response/Short Answer**

- Multiple-choice
- True/False
- Matching
- Fill in the blank (one to a few words)
- Label a diagram

**Written Response**

- Writing in response to a question or request, e.g.,  
 “How \_\_\_\_\_ are these two \_\_\_ similar?” or  
 “Explain the effects of the Stamp Act on the colonists.”

**Performance Assessments**

- Demonstrating skills
- Developing products

**Target-Method Match**

Mark a plus sign (+) in each assessment method column that you believe represents a good match for each kind of learning target.

**Kind of Target**

**Method**

	Selected Response	Extended Written Response	Performance Assessment
Knowledge			
Reasoning			
Performance			
Product			

## Links Between Achievement Targets and Assessment Methods

TARGET TO BE ASSESSED	ASSESSMENT METHODS			
	SELECTED RESPONSE	EXTENDED RESPONSE	PERFORMANCE ASSESSMENT	PERSONAL COMMUNICATION
<b>KNOWLEDGE &amp; UNDERSTANDING</b>	Good match for assessing mastery of elements of knowledge.	Good match for tapping understanding of relationships among elements of knowledge.	<i>Not a good match – too time consuming to cover everything.</i>	Can ask questions, evaluate answers, and infer mastery, but a time-consuming option
<b>REASONING PROFICIENCY</b>	Good match only for some patterns of reasoning	Written descriptions of complex problem solutions can provide a window into reasoning proficiency	Can watch students solve some problems or examine some products and infer about reasoning proficiency	Can ask student to "thing aloud" or can ask follow up questions to probe reasoning
<b>SKILLS</b>	<i>Not a good match. Can assess mastery of understandings prerequisite to skillful performance, but cannot rely on these to tap the skill itself</i>		Good match. Can observe & evaluate skills as they are being performed	Strong match when skill is oral communication proficiency; also can assess mastery of knowledge prerequisite to skillful performance
<b>PRODUCTS</b>	<i>Not a good match. Can only assess mastery of the understandings prerequisite to the ability to create quality products</i>	Strong match when the product is written. Not a good match when the product is not written	Good match. Can assess the attributes of the product itself	<i>Not a good match.</i>

Source: *Classroom Assessment for Student Learning*, p. 100, R. Stiggins, J. Arter, J. Chappuis, and S. Chappuis

©2008, ETS Assessment Training Institute, Portland, OR [www.ets.org/ati](http://www.ets.org/ati) Permission granted for school district use, subject to "Conditions for Approved Use of ETS/ATI Materials" included in this document

## Sample achievement targets across school subjects

<b>Achievement Target</b>	<b>Reading</b>	<b>Writing</b>	<b>Science</b>	<b>Math</b>
Know	Background knowledge required by text Vocabulary	Vocabulary Mechanics of usage Knowledge of topic	Facts and concepts	Number meaning Facts Numeration system
Reason	Process the text & comprehend the meaning	Choose words & syntactic elements to convey message Evaluate text quality	Hypotheses testing Classifying species	Identify & apply algorithms to solve problems
Performance Skills	Oral reading fluency	Letter formation Keyboarding skills	Manipulate lab apparatus correctly	Use manipulatives while solving problems
Products	Diagram revealing comprehension	Sample of original text	Written lab report Science fair model	Diagram depicting problem solving solution

©2008, ETS Assessment Training Institute, Portland, OR [www.ets.org/ati](http://www.ets.org/ati) Permission granted for school district use, subject to “Conditions for Approved Use of ETS/ATI Materials” included in this document.

1. The amount of fat, lean muscle, bone, connective tissue, and water you have is called your

- a. muscle composition
- b. physical fitness
- c. body composition
- d. cardio respiratory fitness

Comment:

2. Metabolic rate

- A. decreases in adulthood
- B. increases in adulthood
- C. stays the same throughout life
- D. decreases in adult men only
- E. increases in adult women only

Comment:

3. The Pol POT regime officially proclaimed that the ethnic minorities totaled only 1 percent off the population; “99 percent” were allegedly Khmers What does word allegedly mean?

- A. Not
- B. Supposedly
- C. Unlikely
- D. Never

Comment:

4. Which is not true about steroids?

- a. causes breast enlargement in men
- b. decreases or causes ovulation to stop in women
- c. increases potency in males
- d. causes "road rage" in both men & women
- e. causes weakening & tearing of ligaments

Comment

5. Eating too much salt may cause which of the following health problems?

- a. liver cancer
- b. high blood pressure
- c. gum disease
- d. don't know

Comment:

Read True/False bottom of page 64 to top page 67 in Using Writing & Evaluating Force Choice Items (Goldenrod)

1. The ONLY sure way to avoid an STD is through abstinence  
a. true                      b. false

Comment:

T

F

2. There are no structural decisions made until phase 6, "long term revisions".

Comment:

T F 3. Tournament badminton players use feathered shuttlecocks because goose feathers are more bulky than hummingbird feathers.

Comment:

T F 4. In 1911 Hiram Bingham, who became a U. S. senator and governor of Connecticut, did not discover the lost Inca city, Machu Picchu, because of his considerable disinterest in antiquity.

Comment:



Read Matching bottom of page 67 to top page 69 in Using Writing & Evaluating Force Choice Items (Goldenrod)

## A

Directions: Match the letters of items in the list at the right with the phrases in the list at the left. Each letter may be used once, more than once, or not at all.

- |  |                           |
|--|---------------------------|
| ___ 1. Formerly ruled by czars         | a. The Great Wall         |
| ___ 2. Site of the Valley of the Kings | b. USSR                   |
| ___ 3. The People's Republic of China  | c. Egypt                  |
| ___ 4. 1776                            | d. A nation is created    |
|  | e. Great Britain          |
|  | f. America is discovered. |

**Comment:**

## B

Match the following terms with the correct definition.

- a. Energy
- b. Input force
- c. Power
- d. Lever
- e. Output force

- 57. A human arm or leg is an example.
- 58. Force you apply to make a simple machine.
- 59. The force a machine applies to accomplish a task
- 60. The ability to do work.
- 61. The amount of work performed over time.

**Comment:**

1. Recent research suggests that \_\_\_\_\_ success is most often due to \_\_\_\_\_

**Comment:**

2. Tyler's purpose for a statement of objectives is to \_\_\_\_\_ so that instructional activities can be planned and developed.

**Comment:**

## Sample Tasks to Analyze

### Example 1: Math Problem

The following problem is one of several given to students to assess problem solving in math. Student work is assessed using a rubric (not included here) that covers mathematical understanding, problem-solving processes, and communication.

A group of 8 people are all going camping for 3 days and need to carry their own water.

They read in a guidebook that 12.5 liters are needed for a party of 5 people for 1 day.

Based on the guidebook, what is the minimum amount of water the 8 people should carry all together? Explain your answer.

*Source:* Used on various assessments in Oregon and Idaho.

### Example 2: Motorized Vehicle—High School

This is a problem in applied physics. It is a culminating project to assess knowledge of motion, written and oral communication proficiency, and reasoning. Design and construct a motorized vehicle that can produce at least two simultaneous motions in different directions to accomplish an action. When your work is done, you will demonstrate your device and explain how it works. In addition, you will be asked why you made certain design decisions, relying on your understanding of physics concepts. Finally, you will also be asked how well your device worked and to explain how you might modify your vehicle to make it better. The attached criteria (here not reproduced, but given to the students) will be used to judge your work.

*Source:* Abstracted from Dorothy Bennett, Assessment and Technology Videotape, Center for Technology in Education, Bank St. College of Education, 610 W. 112th St. New York, NY 10025

### Example 3: Create a Flag—Middle School

The following is a task, as given to students. There is no additional explanation.

As a group, you have been asked to submit a design for the flag for the new Pacific Island nation of Koluhra.

Koluhra is located in the Caroline Islands of Micronesia approximately 400 miles SE of the state of Yap at 3 degrees N. latitude, and 147 degrees E. longitude.

The inhabitants of this new nation have migrated over time from outer islands of three surrounding island nations following a series of typhoons. Traditional chiefs have formed a governing body and share a vision of unity among the people upheld by extended family relationships. They hope to carefully expand their economy around ecotourism.

Using what you know about the islands of the western Pacific, the information provided above and your joint creativity, design a flag that can serve as a visual symbol of Koluhra identity. You will need to prepare to exhibit your design to the council of chiefs and explain how each element of the flag contributes to the identity of Koluhra.

*Source:* Abstracted from Regional Educational Laboratories (1998), Toolkit98, Appendix B, Sample B5, contributed by Kathy Busick.

Source: Classroom Assess for Student Learning – Supplementary Material

©2008, ETS Assessment Training Institute, Portland, OR [www.ets.org/ati](http://www.ets.org/ati) Permission granted for school district use, subject to “Conditions for Approved Use of ETS/ATI Materials” included in this document.

### Example 5: Sow Bugs—Grades 4–6

This task was one of three given to students to assess science process skills.

Students receive five sow bugs, a round dish to contain them, a bright light and strips of dark cardboard to create regions of light and dark, filter paper, a spray bottle for creating damp regions, and a stopwatch. The students are to answer the following questions:

- Do sow bugs prefer light or dark environments?
- Do sow bugs prefer damp or dry environments?
- Do dampness and amount of light in combination make a difference in sow bug preferences?

Scoring is procedure based (task specific). For each experiment, observers focus on the method used to solve the problem, the adequacy with which conditions are manipulated, the measurement strategies used to determine the results, and the correctness of the solution generated.

*Source:* Adapted from “On the Stability of Performance Assessments,” by M. A. Ruiz-Primo, G. P. Baxter, and R. J. Shavelson, 1993, *Journal of Educational Measurement*, 30, pp. 41–51; “Performance Assessments: Political

### Example 6: Fruit—Grade 7

The following is a seventh-grade math problem. Scoring is based on the correctness of the answers and the steps used to solve the problem. Lisa put some fruit in a large bowl for her friends. The bowl had twice as many apples as oranges, and half as many pears as oranges. Altogether, there were 14 pieces of fruit in the bowl.

- How many apples did Lisa put in the bowl?
- How many oranges?
- How many pears?
- Explain or show how you found each answer.

### Example 7: The Car Problem—Post High School

The following problem was given to students to determine their ability to solve problems. It was scored for the correctness of each step.

From the classified section of a newspaper, select one particular brand and model of automobile that appears several times. Collect data on the age (number of years old) of the vehicle versus the asking price. You should have at least 8 points.

- a. Plot the data you have gathered. Carefully label your graph.
- b. Draw an “eyeball fit” line through the data. (Your next task will be easier if the line goes through two of the data points.)
- c. Write an algebraic linear model to describe the line you have drawn.
- d. Interpret the meaning of the slope in your model.
- e. Interpret the meaning of the vertical-intercept of your model. (Include the numerical value and the units.)
- f. In there are other data points that do not seem to fit the overall linear pattern of the other data, try to explain why.

*Source:* Collected from a mathematics teacher at Mt. Hood Community College, approximately 1990.

*Source:* Classroom Assess for Student Learning – Supplementary Material

©2008, ETS Assessment Training Institute, Portland, OR [www.ets.org/ati](http://www.ets.org/ati) Permission granted for school district use, subject to “Conditions for Approved Use of ETS/ATI Materials” included in this document.

### Example 8: Math Portfolio—Grades 4, 8

Students are asked to assemble a portfolio that demonstrates their mathematical ability.

Each portfolio should contain 10-20 selections.

5-7 of these should be “best pieces” and must include: 1 puzzle, 1 investigation, 1 application, and no more than 2 pieces of group work.

- The student can select other pieces that demonstrate ability.
- The student should write a letter to the evaluator that describes what he or she has chosen for his or her portfolio and what it shows about the student.

The portfolios are assessed using rubrics that cover the following:

**Problem Solving:** How well does the student understand the problem, how does the student solve the problem, why does the student solve it the way she or he did, and what observations, connections, generalizations does the student make about the problem?

**Communication:** What terminology, notation, symbols does the student use to communicate his or her math thinking, what representations (graphs, charts, tables, models, diagrams, pictures, manipulatives) does the student use, how clear is the student's communication of mathematical thinking and problem solving?

Source: Abstracted from Vermont State Department of Education, Portfolio Assessment, about 1984.

Source: Classroom Assess for Student Learning – Supplementary Material

©2008, ETS Assessment Training Institute, Portland, OR [www.ets.org/ati](http://www.ets.org/ati) Permission granted for school district use, subject to “Conditions for Approved Use of ETS/ATI Materials” included in this document.

### **Sample 9: Grade 10, Spanish; Task Title: DGS and Our World: A Mural of Our Time**

**Approx. time frame: 2 weeks**

**What desired understandings/content standards will be assessed through this task?**

- 1. Art reflects the culture of its time and changes in society.**
- 2. Political, social and historical events influence the artwork of a period.**
- 3. Artwork conveys the events and feelings of the artist's life.**

**Standards:**

**29B1b: Demonstrate one or more art forms representative of areas where the target language is spoken.**

**29B5a: Explain the cultural and historical significance of characteristic art forms of a target language society. What is the purpose of this assessment task?      summative**

**Performance task:**

**Board of Education of District 99 would like to have a piece of student artwork featured near the entrances of the new areas of both buildings. They have asked each Spanish 3 class to design a mural depicting the lives of DGS students in the early 21<sup>st</sup> century, the time when the buildings were completed. Each class will act as a team of artists working cooperatively to design a mural that depicts the culture and recent social/historical events of the school as well as the culture and recent crucial events in America. Each mural must contain an aspect that is personal to each student in the class.**

**Upon completion of the ideas and a sample, the mural will be presented to the board at a meeting. In the presentation the following must be addressed: 1) An explanation of the mural and its components with a direct tie to how what was learned in the muralist unit helped guide you in the construction of your own mural. 2) An explanation of the process the class went through to arrive at the finished product. 3) A sharing of reflections on what was learned about the content and the process. Those reflections will be drawn out of the written reflections completed by each student. From all proposals the board will select one mural to be constructed at each site, based on the how well the overall task and its explanation were completed.**

**By what criteria will student products/performances be evaluated?**

- How well the mural depicts the culture and recent events of the school and the country.**
- How well the design is tied to information learned in class about artists and their murals.**
- How clear the explanation of the process and product is.**
- Whether or not there is evidence of participation by each student in the entire process based on teacher observations and written reflections.**

**What type of scoring tools will be used for evaluation? \_analytic rubric**

**Sample 10: Grade 8, Task Title:** Children’s storybook **Approx. time frame:** 1 week

**What desired understandings/content standards will be assessed through this task?**

**Enduring understandings (Students will have one of these)**

1. Different perspectives may lead to conflict/discrimination.
2. Scapegoating gives reason for discrimination.
3. Perspective may change with increased industrializing of social, political, and economic conditions.
4. Indifference perpetuates discrimination.
5. Discrimination occurs from a lack of social, political, and economic understanding.

**Content**

16.D.3 (W) Identify the origins and analyze consequences of events that have shaped world social history including famines, migrations, plagues, and slave trading.

2.B.3b Compare and contrast common literary themes across various societies and eras.

2.B.3c Analyze how characters in literature deal with conflict, solve problems and relate to real-life situations.

**Process**

3.B.3a Produce documents that convey a clear understanding and interpretation of ideas and information and display focus, organization, elaboration and coherence.

**What is the purpose of this assessment task?** summative

**Performance task:**

You are a children’s author who wants to retire, but you need to write one last book before your publishing company allows you to retire. The publishing company wants you to write one last book about discrimination because they feel children need to know what discrimination is early on in their life. If the children know about discrimination in their formative years, hopefully they won’t discriminate. Your target audience is children ages 6-10. You will be judged/paid on the appropriate plot that shows the theme of discrimination, appropriate for age group, neatness, content, mechanics, and overall appearance of your manuscript.

**Sample 11: Grade 9 Science/S. Studies, Task Title:** Young Scientist Competition **Approx. time frame:** 1wk

**What desired understandings/content standards will be assessed through this task?**

7.C Select and use appropriate technology, instruments and formulas to solve problems, interpret results and communicate findings.

13.A. Know and apply the accepted practices of science.

13.A.5b Explain criteria that scientist use to evaluate the validity of scientific claims and theories.

16.A.4a Analyze and report historical events to determine cause-and-effect relationships.

**Understandings:**

Opportunity and need interact to initiate discovery and change knowledge.

Scientists evaluate physical phenomena and data to ascertain whether or not theories or claims are valid.

The cause-and-effect relationship between data and discovery results in selecting appropriate tools and interpret the results.

**Process:**

Make generalizations regarding change and discovery

**What is the purpose of this assessment task?** \_\_\_ formative \_\_\_x\_\_\_ summative

**Performance task:**

Students will be given the following scenario:

The Lyons Club is sponsoring a scholarship contest for high school students who will attend college. The deadline for entry is one week from today. The goal of the contest is to design a presentation using technology that depicts a scientific discovery in the last 100 years. You will present your project to the scholarship board of directors comprised of a scientist, a historian, and three members from the organization. You need to develop a presentation explaining the aspects of scientific discovery focusing on the discovery of your choice so that it can be displayed in an exhibit and understood without further explanation from you. A successful

entry will not only explain the discovery, but it will relate it to past discoveries and show what is common to all discoveries. The presentation must also give your interpretation as well as experts in the field of the changes the discovery caused during the time it was initially discovered and the present impact it has. In addition, you must also give plausible applications for the discovery in the future. Finally, a project that will be in the running for the scholarship must also give perspective on the person who made the discovery and what led that person to make the discovery. Good luck, and may the force be with you!

**Sample 12: Grade 9 English, Task Title: “Should We Chuck Huck?” Approx. time frame: 1 week**

**What desired understandings/content standards will be assessed through this task?**

### **Enduring Understandings**

- Moral responsibility requires an individual to act according to his/her conscience, even when the consequences of the act will be unpleasant.
- Financial status, reputation, and ethnicity dictate one’s position in the social hierarchy rather than one’s moral or immoral behavior.

### **Content Standards**

- 1.C.4b: Explain and justify an interpretation of a text.
- 1.C.4d: Summarize and make generalizations from content and relate them to the purpose of the material.
- 2.A.4b: Explain relationships between and among literary elements including character, plot, setting, theme, conflict and resolution and their influence on the effectiveness of the literary piece.
- 2.B.4b: Analyze form, content, purpose and major themes of American literature and literature of other countries in their historical perspectives
- 2.B.4c: Discuss and evaluate motive, resulting behavior and consequences demonstrated in literature

### **Process Standards**

- 3.A.4: Use standard English to edit documents for clarity, subject/verb agreement, adverb and adjective agreement and verb tense; proofread for spelling, capitalization and punctuation; and ensure that documents are formatted in final form for submission and/or publication.
- 5.C.4a: Plan, compose, edit and revise information (e.g., brochures, formal reports, proposals, research summaries, analyses, editorials, articles, overheads, multimedia displays) for presentation to an audience.

**What is the purpose of this assessment task?**      summative

### **Performance task:**

Over the summer, a committee has been formed by a local school board to address concerns voiced by parents in its community about use of stereotypical character depictions and racial slurs in *The Adventures of Huckleberry Finn*, which is currently being taught in the local high school. This committee wants to look into the matter and determine whether or not to include *The Adventures of Huckleberry Finn* in their freshman curriculum. As a literature professor from the University of Illinois, you have read about the controversy and, since you have studied the novel extensively, feel that your insight would be helpful to the committee. You have contacted the board of education, and they have requested that you write a letter expressing your opinion on the issue to help them make an informed decision regarding the novel’s inclusion in the local

curriculum. You must decide if the book is appropriate and valid to teach to freshmen today. Your goal is to persuade the school board to include or not include the novel in the curriculum. The parents are concerned that the book's racial overtones overshadow any literary merit the book may have, making it inappropriate for a young, impressionable audience. Therefore, the school board has requested that if you feel the book should be taught, your letter should be focused around a generalization that includes two timeless concepts, giving the book modern-day credence. If, however, you feel that the racial overtones overshadow the book, you must justify why the overtones are more pervasive than the most far-reaching generalization. To support your generalization, you must use at least two relevant examples and supporting quotations from the novel. It is also vital that your letter be taken seriously by the board; therefore, you must follow authentic letter format and use standard grammatical conventions.

**By what criteria will student products/performances be evaluated?**

- Is the letter focused around a generalization that includes two concepts?
- How well is the generalization supported by relevant examples and quotes from the text?
- How well does the letter follow authentic letter format?
- How well does the letter follow standard grammatical conventions?

**How will students self-assess?**

Students will edit a peer's letter, complete the assignment rubric, and reflect on their effort and performance on a teacher questionnaire.

**What type of scoring tools will be used for evaluation?** analytic rubric

**Sample 13: Grade 7 math/English, Task Title: *Peanut Butter Commercial*** Approx. time frame: 2 class periods

**What desired understandings/content standards will be assessed through this task?**

- ✓ Data creates change.
- ✓ Data impacts the media.
- ✓ Media uses data to persuade society for positive or negative change.
- ✓ Media can use data to create biases that misinform the public.
- ✓ Media can use data from their viewpoint causing change that benefits them.

**Content**

- 10A Organize, describe, and make predictions from existing data.
  - 10A.3a Construct, read and interpret tables, graphs (including circle graphs) and charts to organize and represent data.
  - 10A.3b Compare the mean, median, mode and range, with and without the use of technology.
  - 10A.3c Test the reasonableness of an argument based on data and communicate their findings.

**Process**

- 4B Speak effectively using language appropriate to the situation and audience.
  - 4B.3a Deliver planned oral presentations, using language and vocabulary appropriate to the purpose, message and audience; provide details and supporting information that clarify main ideas; and use visual aids and contemporary technology as support

**What is the purpose of this assessment task?** formative



**Performance task:**

Your team (group of three or four) is a group that works in the marketing department of a peanut butter company. The company you work for needs a good commercial to stay in business. Using the data about peanut butter from consumer reports, your group must make a commercial that promotes one of the peanut butters' on the list. Using at least two of the attributes, (quality rating, price per serving, natural, regular, creamy, chunky, salted, unsalted) you must create and use a visual aid in a commercial that demonstrates an understanding of box-and-whisker plots, mean, median, mode, and range. You will present your commercial to the marketing executive so he can decide which commercial to use to keep the business successful. Your commercial must be persuasive in nature and use math concepts learned in class.

**What student products/performances will provide evidence of desired understandings?**

Student commercials (taped or live)

**By what criteria will student products/performances be evaluated?**

- Proper use of math concepts (box-and-whisker plots, mean, median, mode, and range)
- Persuasive Creative

**What type of scoring tools will be used for evaluation?** holistic rubric

# Rubrics: Choices to consider

## 1. Holistic or analytic trait rubric?

- **Holistic** rubric gives a \_\_\_\_\_ score or rating for an entire product or performance based on an overall perception of the student's work.
  - When to use
    - Judging simple products or performances ( don't have more than one trait i.e. reading fluency or response to essay question
    - Getting a quick snapshot of overall quality or achievement
    - Judging the impact of a product or performance
    - When the speed of scoring is more important than knowing how to precisely describe quality.
  - Disadvantages:
    - Two students can get the same score for vastly different reasons
    - Not as good for identifying strengths and weakness and planning instruction
    - Does not give students feedback to guide their improvement
- **Analytical** trait rubric \_\_\_\_\_ a product or performance into essential traits or dimensions so that they can be judged separately.
  - When to use
    - Planning instruction – show relative strengths and weaknesses
    - Teaching students the nature of a quality product or performance - they need the details
    - Detailed feedback to students or parents
    - When knowing how to precisely describe quality is more important than speed
    - Complicated skills, products, or performance dimensions are needed in order to be clear
  - Disadvantages
    - Scoring is slower
    - Tasks longer to learn

## 2. Generic or Task specific rubric?

- Generic rubric can be used across similar performances i.e. all open-ended math problems, all oral presentations
  - When to use
    - When the rubric is being used instructionally to help students understand the nature of quality and generalize from one task to the next.
    - When students will not be doing exactly the same task; when students have a choice as to what evidence will be chosen to show competence on a particular skill or product.
    - When teachers are trying to judge consistently in different classes or grades
  - Disadvantages

- Takes longer to learn, but therein is also the strength - learning the rubric is also learning the skill
- Takes longer to score, when the only goal is to assess quickly
- Sometimes, task specific features have to be added if particular knowledge is desired.
- Task specific can only be used for a single task
  - When to use
    - When speed of getting a score is more important than thinking through what is being scored.
    - When you want to know if students know particular facts, equation, methods, or procedures
    - When consistency of scoring is of utmost importance
  - Disadvantages
    - Can't show to student ahead of time because it "gives away" the answer
    - Have to develop a new rubric for each task; this takes time and sometimes this isn't even possible
    - Does not make the rater think – scoring is on automatic pilot
    - "Right" answers not in the scoring guide are sometime missed
    - Does not help define the nature of quality in general; only states what quality looks like for a particular task

### **3. How does one decide on a rubric scale?**

- Generally, the more open ended and complex the performance, the broader the scale range
- Have enough score points to distinguish quality (fewer than 4 is not enough & more than 8 too many)
- If primary purpose is to certify competence to a standard, then 4 point scale, in which 3 is competent.
- A developmental rubric used across grade level to track progress continuously from grade to grade, a larger scale is needed.
- Have a good reason for the scale you use. Some people avoid odd number of points because tendency to go to the middle.
- Some teachers stay away from 5 point scale because looks like typical A-F grading scale.
- Number of points depends on the purpose of the rubric and the nature of what is to be assessed.

Source: Scoring Rubrics in the Classroom, J Arter, J. McTighe, 2001.

## When to Use Different Kinds of Rubrics

	Large Scale*	Grading	Teacher uses: Plan instruction, Feedback to students	Student uses: Self-Assessment, Track Progress, Communicate
Knowledge-Specific knowledge to be learned by all students	Task specific	Task specific	Task specific	Task specific
Knowledge-Student choice as to knowledge to be demonstrated	Probably never happen	General holistic	General holistic	General holistic
Reasoning proficiencies, Performance skills, Products	Task specific or general holistic	Task specific or general holistic unless the grading rubric is also used instructionally	General Analytic	General Analytic

\* Large scale assessment takes place across classrooms using the same assessment materials at roughly the same time and under the same conditions, such as benchmark tests, state tests.

Source: *Creating & Recognizing Quality Rubrics*. J. Arter, J. Chappius, ETS/ATI, p.17.

©2008, ETS Assessment Training Institute, Portland, OR [www.ets.org/ati](http://www.ets.org/ati) Permission granted for school district use, subject to “Conditions for Approved Use of ETS/ATI Materials” included in this document.

# How to Develop a General Rubric

**Step 1: Choose a learning target worth the time.**

**Step 2: Search out existing relevant scoring guides/rubrics.**

**Step 3: Gather samples of student work**

- Try to have 20 samples representing more than one topic or task
- A variety of samples helps ensure that all important general criteria ends up on the final rubric.

**Step 4: Sort student work**

- Start with three stacks/categories –strong/proficient, medium/developing, weak/beginning
- Write down the reasons for placing student work in a category.
- Reason need to be descriptive
- For each sample, write down exactly what you are saying to yourself as you place the work in the category. Don't wait until the end.
- Write comments you would make to provide feedback to the student on what was done well and how to make the performance or product better next time.
- Avoid broad statement.....be specific
- As sorting progress, individuals may feel the need to create more or less stacks/categories.
- This is the beginning of determining the final number of levels.

**Step 5: Group like indicators together**

- This is the initial list of separate criteria for the categories/levels.
- At this stage criterion refining takes place.
- Most rubrics will go through several stages of criteria definition and organization.
- The draft is the descriptors from the brainstormed listed sorted into categories.

**Step 6: Identify student work that illustrates each level on each criterion.**

- Start with the extremes – strong & weak
- Find examples for the middle if you are using an odd number of levels.
- Find several different examples that illustrate each level. Find examples across assignments.
- Watch for particular examples of the errors you student commonly make.

**Step 7: Test the rubric/scoring guide and revise it as needed.**

- Possible problems:
  - Student performance or products don't match the descriptors in the rubric.
  - Features of the student work seem to be rated in more than one criterion.
  - Criteria should be subdivided - cover too much or include too many important indicators.
  - Internal structures of the descriptors need to be improved.
  - Content of some levels is not parallel.

**Step 8: Repeat the cycle of scoring and revising.**

*Source: Creating & Recognizing Quality Rubrics. J. Arter, J. Chappuis. ETS, p. 68-80.*

©2008, ETS Assessment Training Institute, Portland, OR [www.ets.org/ati](http://www.ets.org/ati) Permission granted for school district use, subject to "Conditions for Approved Use of ETS/ATI Materials" included in this document.

## Oral Presentation Rubric (Grades 4-12) Criterion I: Content

Strong	Middle	Weak
Ideas are focused and supported with relevant details and examples, Content is relevant for the task.	The topic is fairly broad, but is focused on relevant content. Support is attempted, but doesn't go far enough.	There is little controlling idea, the speaker is still in search of a topic, or the length is not adequate for development.
Information is accurate.	Ideas are reasonably clear; but there are some problems with accuracy.	Information is limited, unclear, or incorrect.
The speaker has chosen the most significant information and stays with the topic.	The speaker generally stays on the topic, but doesn't develop a clear theme,	Everything seems as important as everything else,
The speaker anticipates the information needs of the audience, adapts content to the listeners' background, and/or refers to listeners' experience,	The listener is left with questions, There seem to be some "holes" in the information,	The presentation may be repetitious or sound like a collection of disconnected thoughts,

## Criterion 2: Organization

Strong	Middle	Weak
The speaker helps the listener understand the sequence of ideas through organizational aids such as previewing the organization, using transitions, and summarizing. Listeners can put the ideas in an outline.	The sequence and relationships are fairly easy to follow, but sometimes you have to make assumptions to connect the ideas. An outline of the ideas requires inferences,	Ideas that go together are not put together: Listeners would have trouble putting the ideas into an outline,
The opening draws the listener in; the closing leaves a sense of closure and resolution,	The presentation has a recognizable opening and closing, but there is little sense of anticipation or closure,	There is no opening or closing,
Details seem to fit where they're placed,	Sequencing is sometimes awkward, Some details don't seem to fit where they're placed,	Sequencing is confusing,

## Criterion 3: Delivery

Strong	Middle	Weak
Volume is loud enough to be heard and understood. Volume is intentionally used to keep the listener's attention and/or enhance the points being made.	The speaker can be heard and volume doesn't distract the listener; but neither does volume draw attention to important points,	The speaker can't be heard and/or changes in volume distract the listener from understanding the points being made,
Visual aids are used effectively to support and enhance meaning,	Visual aids, while understandable, don't add much to the presentation,	Visual aids are confusing, do not relate to the point being made, or distract the listener:
Pronunciation and enunciation are clear enough to be understood and are used to emphasize important points,	Pronunciation and/or enunciation are generally clear enough to be understood, but are not used effectively to underscore important points,	Pronunciation and/or enunciation detract from being able to understand the speaker:
The speaker exhibits very few disfluencies, such as "ah," "um," and "you know."	While the speaker exhibits disfluencies, they don't detract from the presentation enough to interfere with meaning,	Disfluencies, such as "um," "ah," and "you know," detract from understanding what is being said,
There is little in the presenter's demeanor; dress, or mannerisms that distracts the listener from the message	The presenter's demeanor; dress, or mannerisms sometimes distract the listener; but meaning is not disrupted,	The presenter's demeanor; dress, or mannerisms distract the listener to the extent that meaning is disrupted,
Pacing is right for the audience, The speaker knows when to slow down and when to speed up,	Pacing is fairly good, but at times the speaker goes too fast or too slow for the listeners to keep up,	Pacing is awkward, The listener wants the speaker to either get on with it or not go so quickly,
Sentences are varied and easy to listen to and understand. They attract and hold attention,	Sentences are usually correct and can be understood, but generally lack the flair that maintains attention,	Sentences either ramble, are choppy, or are awkward, Sentence structure might all be the same and so become boring,

Source: Creating & Recognizing Quality Rubrics Copyright © 2006 ETS ©2008, ETS Assessment Training Institute, Portland, OR [www.ets.org/ati](http://www.ets.org/ati)  
 Permission granted for school district use, subject to "Conditions for Approved Use of ETS/ATI Materials" included in this document.

## GENERAL CONCEPTUAL UNDERSTANDING RUBRIC

High	Medium	Low
<p>A high score in conceptual understanding means that the student shows an accurate and extensive understanding of the topic. This can be shown in many ways, including the following:</p> <ul style="list-style-type: none"> <li>• Correct and precise use of terminology.</li> <li>• Precise selection of the pieces of information required to make a point (no more, no less).</li> <li>• Correct and appropriate use of examples and counterexamples.</li> <li>• Few errors in information.</li> <li>• Connections made to other, related topics.</li> <li>• Demonstration of which distinctions are important to make.</li> <li>• Key concepts identified and addressed.</li> <li>• A relevant focus sustained throughout the work.</li> <li>• Relevant use of a diagram or graph; knows when such things will aid understanding.</li> </ul> <p>Concise explanations that are to the point.</p>	<p>A medium score in conceptual understanding means that the student presents some important information, but there is a sense that the student is only about halfway home in terms of understanding. Performance is indicated by the following:</p> <ul style="list-style-type: none"> <li>• Reasonably clear ideas, but the reader needs to make some guesses as to what the student meant.</li> <li>• Even though a general point is made, the student hasn't fine tuned the topic.</li> <li>• Some parts of the work seem repetitive.</li> <li>• The balance in the work seems a little off.</li> <li>• Some vocabulary is used well, some is not.</li> <li>• Some examples and graphics are appropriate, some aren't.</li> <li>• Sometimes the student seems to know which concepts and points are most important and telling; other times not.</li> <li>• Information seems to be based on retelling rather than the student making his or her own connections.</li> <li>• The focus tends to shift.</li> </ul>	<p>A low score in conceptual understanding indicates that the student is still searching for the connections that will make the content meaningful. Weak performance is indicated by such things as the following:</p> <ul style="list-style-type: none"> <li>• Ideas are extremely limited or hard to understand, even when the reader tries to draw inferences based on what is there.</li> <li>• The text may be repetitious or read like a collection of disconnected, random thoughts.</li> <li>• Information is inaccurate.</li> <li>• Terminology is used incorrectly.</li> <li>• There is little sense of which information is most important.</li> <li>• Visual displays, when used, are not helpful or unrelated to any points the student is trying to make.</li> </ul>
<p>Source: Creating &amp; Recognizing Quality Rubrics Copyright © 2006 ETS</p>		

## Scoring Guide for Science Journal

<i>Novice</i>	<p>Writing <i>is messy</i> and contains spelling errors.</p> <p>Pages are out of order or missing,</p>
<i>Apprentice</i>	<p>Entries are incomplete.</p> <p>There might be some spelling or grammar errors,</p>
<i>Master</i>	<p>Entries contain most of the required elements and are clearly written.</p>
<i>Expert</i>	<p>Entries are creatively written.</p> <p>Procedures and results are clearly explained.</p> <p>Journal is well organized; presented in a duotang,</p>

Source: From "What's Still Wrong with Rubrics: Focusing on the Consistency of Performance Criteria Across Scale Levels," by R. Tierney & M. Simon, 2004, *Practical Assessment, Research & Evaluation*, 9(2), p. 5. Used with permission. Creating & Recognizing Quality Rubrics Copyright © 2006 ETS

Source: Creating & Recognizing Quality Rubrics Copyright © 2006 ETS ©2008, ETS Assessment Training Institute, Portland, OR [www.ets.org/ati](http://www.ets.org/ati)  
 Permission granted for school district use, subject to "Conditions for Approved Use of ETS/ATI Materials" included in this document.



## **Devising Extended Response Items**

We have all experienced "essay" questions on tests, some of which may have been painful. "Discuss photosynthesis." "Analyze *King Lear*." "Explain the causes of the Civil War." High-quality extended written response items, in contrast to these, carefully frame the task so that students who have learned the material know how to tackle it.

## **ITEMS ASSESSING KNOWLEDGE MASTERY**

Extended written response items that assess factual and conceptual knowledge do three things: (1) set a clear and specific context, (2) indicate what students are to describe or explain, and (3) point the way to an appropriate response without giving away the answer. To assess the learning target "Understand the importance of the carbon cycle and how it works," we could create this item:

### *1. Set the Context*

We have been studying the importance of the carbon cycle and how it works.

In this example, the context is stated as a paraphrase of the learning target. With extended written response questions, especially if they are included in a test with other items, it helps students to be reminded of the specific body of knowledge they are to use when framing their responses.

### *2. Tell What to Describe or Explain*

Based on your understanding of the carbon cycle, describe why we need to know about it and how it works.

The task they are to carry out is stated in the second part-with a knowledge or conceptual understanding learning target, students are generally explaining or describing something. Notice this sentence doesn't just say "Describe the carbon cycle," for two reasons. First, that would not be sufficient to assess the learning target, and second, it doesn't give enough guidance-describe what? If you want students to describe how it works, make sure the item lays out that expectation.

### *3. Point the Way to an Appropriate Response* Be sure to include the following:

- Why it is important to understand the carbon cycle (5 points)
- The four major reservoirs where carbon is stored (4 points)
- At least six ways that carbon gets transferred from one place to another (6 points)

The third part helps students know what will be considered an appropriate and complete response; students who know the material will be able to answer well, and students who don't know it won't be able to bluff their way through.

## ITEMS COMBINING KNOWLEDGE MASTERY WITH REASONING

Extended written response items that ask students to reason with knowledge they have learned are similar in structure to items that assess knowledge mastery. They also have three components, slightly altered to account for the addition of reasoning: (1) set a clear and specific context; (2) specify the kind of reasoning to be used; and (3) point the way to an appropriate response without giving away the answer.

For example, to assess the sixth-grade learning target, "Explain how an author develops the point of view of the narrator or speaker in a text," (CCSSI 2010a, p. 36), the Common Core State Standards document offers this extended written response task as a possibility: "Students explain how Sandra Cisneros's choice of words develops the point of view of the young speaker in her story 'Eleven'" (CCSSI, 2010b, p. 89).

Applying our item frame to the task it might look like this:

### *1. Set the Context*

We have been studying "point of view"-what it means and how to identify it in a story.

### *2. Describe the Reasoning Task*

After reading the story "Eleven," explain how the author uses word choice to show the point of view of the young speaker.

### *3. Point the Way to an Appropriate Response*

Choose at least three examples. Make sure you explain what the young speaker's perspective is and how each example shows that.

(Written response items should also include information about how they will be evaluated, whether with a list of a number of points or with a rubric.)

As another example, to assess the learning target, "Evaluate opposing positions on humankind's role in global climate change," we could create this item:

### *1. Set the Context*

There are those who contend that global climate change is a naturally occurring phenomenon and others who contend it is caused by the actions of humans.

### *2. Describe the Reasoning Task*

Analyze the evidence we have studied to support each claim. Decide whom you think has the stronger argument. Defend your judgment with reasons.

### *3. Point the Way to an Appropriate Response*

In doing so, consider the evidence from geologic history, the history and levels of emissions, and the political and economic interests of each side.

Here are two simpler examples:

- Explain the mathematics formula we studied today in a memo to a student who was absent.
- Teach younger students how to read a contour map by creating a list of instructions accompanied by diagrams and/ or illustrations.

Interpretive items allow you to assess mastery of specific patterns of reasoning disentangled from student mastery of the prerequisite content knowledge. As you will recall from Chapter 4 we do this

by supplying a passage, table, chart, or map of background information about a given topic and then asking students to write a response demonstrating the targeted pattern of reasoning: describe certain relationships, draw comparisons, conduct analyses, or create and fill categories, for example.

Interpretive items also do three things: (1) set the context, (2) describe the reasoning task, and (3) point the way to an appropriate response without "over helping."

To assess the learning target "Summarize text," we could create this item:

*1. Set the Context*

We have been learning how to write a summary—a brief statement of the main ideas of a text.

*2. Describe the Reasoning Task*

After reading (the assignment provided), write a paragraph that summarizes the main ideas.

*3. Point the Way to an Appropriate Response*

**In your paragraph, be sure to do the following:**

- focus only on the main ideas (2 points)
- include enough information to cover all of the main ideas (2 points) *or*

Your paragraph will be evaluated with the Summary Rubric, attach

Source: Creating & Recognizing Quality Rubrics Copyright © 2006

ETS ©2008, ETS Assessment Training Institute, Portland, OR [www.ets.org/ati](http://www.ets.org/ati) Permission granted for school district use, subject to "Conditions for Approved Use of ETS/ATI Materials" included in this document.

## Written Response Items to Analyze

1. In your opinion, which scene in *The Diary of Anne Frank* forms the dramatic climax of the play? Does the course of events or the main character's course of action change at this point? Is the main conflict resolved at this climax, or do further complications arise?
2. The fourth grade students have been studying the Holocaust, and want to find out more about the way the Nazi invasion into European countries affected the lives of the boys and girls who lived in those countries. Your task is to take the role and perspective of a different character each day from your readings in Number the Stars and explain their point of view in reaction to the events that happened to them, which causes a change in their life as a result of the Nazi invasion. At the end of the 2 weeks, you will be partnered with a fourth grader to share your journal entries and have a discussion on the impact on the characters' lives as a result of the Nazi occupation. Your journal will be judged for the inclusion of five entries for each character showing their point of view. In addition, you will be evaluated on your presentation's expressive and well-paced delivery, as well as clarity and eye contact. From this experience, you will gain an understanding of how one's point of view is affected by both coping with the change and striving for continuity in one's life situation.
3. Imagine a world where cars do not exist. In this world not only do cars not exist, but one that would be very confused even scared of the accelerator because they do not know what it means to accelerate. Even more basic than acceleration, they do not know what makes something move. (Have them begin to describe this world in groups. Once they begin to discuss this world in their groups), this world also does not have any idea what baseball is, and if it would not make any sense to them that if a batter hit a ball hard enough the bat would vibrate in their hands. This action might even be considered witchcraft and bring about a terrible plague upon the people. (Have the students add these elements to the world in which they are making describing. Once they have successfully determined the world in which all of the previous descriptions apply,) You are a scientist that lives in this world. Being the smart and successful scientist that you are, you know and understand what acceleration is and what causes it. You also know why the bat vibrates. It is your job to convince these people in the next two weeks your discovery of acceleration, change in motion, and what makes the baseball bat vibrate. Since these people are *very* religious, this is not going to be an easy task, since they remember the Anglican Church and the government are one entity and therefore any new ideas that might challenge the church could lead to your imprisonment. Recall it was just a few hundred years ago that Galileo was imprisoned for his view of the universe. In order to convince the people and the church of your discoveries, you must provide experimental evidence, examples of how these ideas are present in the everyday life and be able to explain them to people who have never heard of the ideas.
4. Satire is a tool that Mark Twain uses to try to open people's eyes to the need for change by exposing the flaws of a person or society. Satirists' main weapon is humor. Discuss one humorous or exaggerated situation Huck encounters that Twain uses to expose human or societal flaws.  
**By what criteria will student responses be evaluated?** A tool for each phase of the design process included in the portfolio with a reflection.  
5: Response shows a masterful understanding and application of the literary term by clearly linking the term to a character or situation from the book and then developing the response thoroughly with in-depth analysis.  
3: Response shows a limited understanding and application of the literary term by partially linking the term to a character or situation from the book and then incompletely developing the response with mediocre analysis.  
1: Response shows a flawed understanding and application of the literary term by failing to correctly link the term to a character or situation from the book and neglecting to develop the response with analysis.

5. Advanced Algebra Ch. 5 Test

#1 In this chapter we discussed three different methods that could be used to solve a quadratic equation. Please list three methods and discuss the strengths and weaknesses of each solving method.

#2 Describe the possible outcomes when solving a quadratic equation. Please provide a sketch of the graph(s) of the corresponding function(s) to support your description of each outcome.

Excerpt from Classroom Assessments for Student Learning: Doing It Right – Doing It We3ll by J. Chappius, R. Stiggins, S. Chappius, J. Arter; Chapter 6: Written Response, p. 181-191.

## PREPARING THE SCORING GUIDES

Again not surprisingly, a key to successful use of written response assessment is having clear and appropriate criteria by which to judge the quality of student responses. We recommend that you don't use "floating standards," in which the evaluator waits to see what responses come in before deciding what will count and thus how to score. Floating standards destroy the validity and reliability of the assessment. Teachers and students alike need to be clear in advance regarding which aspects of the response are important-this is as essential to a quality assessment as is thoughtful item development. In this section, we explain three scoring guide options and describe how to create each.

### Scoring Guide Options

A *scoring guide* is simply a way to assign points to specific features of a response. Three types of scoring guides are appropriate for assessing written response items: lists, task-specific rubrics, and general rubrics.

**LISTS.** A scoring guide in list form identifies the possible correct answers or desired features of the response and specifies how points will be awarded. Use a list when the desired answer has several parts and each one represents a specific category of knowledge or reasoning with knowledge.

For an item asking students to cite instances where Spanish literature and politics may have influenced each other in the twentieth century, the scoring guide might look like this:

3 points for each instance cited, maximum 9 points

Quality of inferences about prominent novelists

Quality of inferences about political satirists

Quality of inferences about prominent political figures of Spain

3. For an item asking students to explain the Krebs cycle, the scoring guide might look like this:

One point for each of the following, maximum five points:

Cycle describes the sequence of reactions by which cells generate energy

Cycle takes place in the mitochondria

Cycle consumes oxygen

Cycle produces carbon dioxide and water as waste products Cycle converts

ADP to energy-rich ATP

For an item requiring students to explain how carbon moves from one place to another, the scoring guide might look like this:

One point for any six of the following, maximum six points:

Carbon moves from the atmosphere to plants through photosynthesis. Carbon moves from the atmosphere to oceans by dissolving in places it is cold.

Carbon moves from the oceans to the atmosphere by evaporation where it is hot.

Carbon moves from land to the atmosphere through fires.

Carbon moves from land to the atmosphere through volcanic eruptions. Carbon moves from land to the atmosphere through burning fossil fuels. Carbon moves from the land into the oceans through erosion.

Carbon moves from plants/ animals to the ground/sediments through decay.

**Creating Lists.** Because a scoring list is simply a description of specific information along with its point value, it is fairly easy to create one. Notice in the last two examples the correct responses are propositions, which we explained in Chapter 5. One good way to create a list of correct responses is to follow the instructions for proposition development found in that chapter. Another form of list useful for scoring is one that describes the characteristics of a correct response. For Example 6.3 shows a mathematics extended written response item scored by a task-specific list following this pattern.

**Sharing Lists.** Because the item should communicate to students what is most important to attend to in a quality response, it should include information about how their work will be scored. In the first "Spanish literature and politics" example, we suggest including the scoring list as written. However, in the second and third examples, the scoring list is the answer key, so we would not recommend including it in the item. Instead, for the second "Krebs cycle" example, the item might conclude with the statement, "5 points possible, one for each key feature described correctly." For the third "carbon cycle" example it might look like this: "6 points possible, one for each correct explanation."

**RUBRICS.** A scoring guide in the form of a *rubric* is a detailed description of the features of work that constitute quality. In a rubric, the features of quality are described at different levels, representing a continuum of "novice" to "proficient" or "weak" to "strong." The levels can be labeled with words, symbols, or numbers. With a rubric, the level is the score.

There are two basic forms of rubrics: *task-specific* and *general*. A *task-specific* rubric describes the features of quality as they appear in a single item or task. For Example 6.4 shows a rubric used to evaluate a task given to assess the learning target "Interpret data from a graph." It can only be used to score one item, so it is called a *task-specific* rubric. A *general* rubric describes the features of quality as they apply across items or tasks. Example 6.5 shows a rubric that can be used to score any item that tests students' ability to interpret information from a graph. It is *general* because it is *generalizable* across tasks. We recommend the use of general rubrics over task-specific rubrics whenever practical for several reasons:

Task-specific rubrics can't be handed out to students in advance because they give away the "answer."

You have to create a new one for each assignment.

With task-specific rubrics, it's easy to award points to features that are idiosyncratic to the specific assignment and not essential to the accomplishment of the learning target.

However, there are times when a task-specific rubric makes sense to use.

## Creating Task Specific Rubrics

Task-specific rubrics are appropriate for assessing conceptual understanding. An example of a second-grade learning target for which we might create a task-specific rubric is, "Understand that the Earth's rotation causes night and day."

To write a task-specific rubric, we refer back to selected response methodology for the "bones." First we create a proposition—a sentence that accurately states the conceptual understanding—and from that we identify statements representing partial understanding, misunderstanding, and lack of understanding. These become our rubric levels.

Here is the process illustrated with the rubric for the item "Day and Night" shown in For Example 6.6:

1. Create a proposition—a sentence that accurately states the conceptual understanding.

### **Proposition:**

"Night and day happen because the earth turns so that the same side is not always facing the sun."

2. Identify typical characteristics of partial understanding. You can also include misconceptions that aren't egregious; that is, they don't contradict the central understanding.

### **Statements of partial understanding:**

"Night and day happen because the moon and sun are on different sides of the Earth."

"The Earth rotates facing the sun and then the moon."

(These are partial understandings because the Earth's rotation does cause day and night, but facing the moon is not a factor. The misunderstanding does not contradict the explanation of Earth's rotation being the central cause.)

3. Identify typical characteristics of misunderstanding or lack of understanding. Also identify any egregious misconceptions that contradict the central understanding, which we call "fatal flaws."

### **Statement of misunderstanding:**

"Night and day happen because the sun moves across the sky." (This is also the "fatal flaw.")

4. Determine how many levels the rubric will have.

Levels:

"2," "1," and "0"

For many task-specific rubrics, it makes sense to have three levels. You can use this formula to create one:

2 points = The response indicates \_\_\_\_\_ *insert statement(s) showing complete understanding*

1 point = The response indicates \_\_\_\_\_ *insert statement(s) showing partial understanding*

0 points = The response indicates \_\_\_\_\_ *insert statement(s) showing lack of understanding, complete misunderstanding, or partial understanding with a 'fatal flaw'*



A three-level ("2," "1", and "0") rubric is a formula for awarding two, one, or no points. It can be used for a short answer or an extended response item, if the conceptual understanding is not too complex.

For conceptual understanding with more variables, it may make sense to have four levels, following this formula:

3 points = The response indicates \_\_\_\_\_ *insert statement(s) showing complete understanding.*

2 points = The response indicates \_\_\_\_\_ *insert statement(s) showing partial understanding. You may also include a statement of simple misunderstanding at this level.* 1 point = The response indicates \_\_\_\_\_ *insert statement(s) showing partial understanding with some misunderstandings, but no "fatal flaw."*

0 points = The response indicates \_\_\_\_\_ *insert statement(s) showing lack of understanding, misunderstanding, or partial understanding with the inclusion of a 'fatal flaw.'* "

A four-level ("3," "2," "1" and "0") rubric is a formula for awarding three, two, one, or no points. It is best used with extended written response items. However, if you find yourself stretched to differentiate clearly between the two-point level and the one-point level, you may want to combine them into a three-level ("2," "1", and "0") rubric.

### **Creating General Rubrics**

Although we often use task-specific scoring, either in the form of lists or rubrics, for assessing student understanding of content, it is not our only option. We can also use a general rubric designed to assign points to student understanding of any concepts within a body of knowledge.

To create a general rubric for conceptual understanding, we can use the task-specific rubric formula itself. Instead of including the content-related statements of understanding, partial understanding, misunderstanding, or lack of understanding, we leave the descriptions in general terms:

2 points = Evidence shows complete understanding

1 point = Evidence shows partial understanding, with no serious misunderstandings

0 points = Evidence shows lack of understanding, complete misunderstanding, or partial understanding with a "fatal flaw"

Or

3 points = Evidence shows complete understanding

2 points = Evidence shows partial understanding with few, if any, simple misunderstandings

1 point = Evidence shows partial understanding with some misunderstanding, but no "fatal flaw"

0 points = Evidence shows lack of understanding, misunderstanding, or partial understanding with the inclusion of a "fatal flaw"

If we are using interpretive items to assess reasoning alone, then all we need is a general rubric for the pattern of reasoning we are assessing.

The situation is different if we're also scoring how well students reason with the knowledge given in the item. In this case, we need both a way to assess content knowledge and a way to assess the pattern of reasoning. For example, if the extended written response item calls for students to make a generalization

based on content, you might use a task-specific list to assess content understanding, and then a general rubric to analyze the quality of the generalization.

We recommend general rubrics for assessing the quality of various patterns of reasoning because they are versatile:

- You can give them to students in advance to help them understand what high quality looks like.
- You can use them to give students feedback on practice work, highlighting strengths and areas for improvement.
- Students can self-assess with them while practicing, evaluating their own strengths and areas for improvement.
- Students can offer peer feedback with them.
- You can use them again and again, for any task calling for the pattern of reasoning.

ETS ©2008, ETS Assessment Training Institute, Portland, OR [www.ets.org/ati](http://www.ets.org/ati) Permission granted for school district use, subject to “Conditions for Approved Use of ETS/ATI Materials” included in this document