Constructed Response

Items
Thought Questions:

- Did this chapter make essays more or less appealing to you after reading it?
- Currently, is assessment using constructed response assessment under utilized, over utilized, or about right on?
Something to Keep in Mind!

- Not everything can be assessed with #2 pencils!

Scientists work to discover how smart whales really are!
Interpretive Exercises

An interpretive exercise consists of a set of data or information, which we'll call a display, followed by a series of problems or questions having answers that are dependent upon the information given.
Interpretive Exercises

Displays may include:

- Prose paragraphs
- Numerical data
- Charts, graphs, diagrams, or maps
- Pictures, drawings, or photographs
- Cartoons or caricatures
- Lists of words or symbols
- Mathematical formulas
- Musical scores or excerpts
- Audio or video recordings
- Poems, short stories, or essays
- Articles from newspapers, magazines, or journals
- Quotations, adages, or scriptures
- Specimens (rocks, plants, animals, chemicals, art, etc.)
One of the methods formerly used by geologists to determine the age of the earth was a calculation based on the amount of salt (NaCl) in the ocean, and the amount added to ocean waters each year by the rivers that empty into the ocean. If this method of age determination is used, certain assumptions must be made. Items 56-62 consist of a number of assumptions. Classify each assumption as:

a. Necessary for the calculation and is probably true.
b. Necessary for the calculation but is probably false.
c. Not necessary for the calculation but is probably true.
d. Not necessary for the calculation and is probably false.

56. The salt concentrations of the oceans is gradually increasing. (a)
57. Oceans have been on the earth since our planet was formed. (b)
58. Ever since its origin, the earth has revolved around the sun. (c)
59. The oceans now contain all the salt that has ever been added to them. (b)
60. The salts which rivers have carried to the oceans have all occurred in mineral form in the rocks before they were dissolved by the river water. (d)
Directions: The following statements refer to the data in the chart above. Read each statement, and mark your answer according to the following key.

T -- if the data in the chart are sufficient to make the statement true.
F -- if the data in the chart are sufficient to make the statement false.
I -- if the data in the chart are insufficient to determine whether the statement is true or false.
1. SAT   **.59**  **.33**  .16  .28  .25  **-.61**  **.55**  **-.37**
2. ACT    **.82**  .03  .31  .35  **.52**  **-.93**  **.60**  **-.62**
3. Grad. Rate  **-.12**  **.36**  .07  .20  **-.64**  .24  **-.68**
4. S/T Ratio  **-.31**  .08  -.36  .15  .20  .27
5. Experience  **.63**  **.68**  **-.39**  .33  **-.29**
6. Salary  **.84**  **-.41**  **.71**  .12
7. Per Pupil  **-.48**  **.59**  **-.06**
8. % Lunches  **-.59**  **.51**
9. Income
10. Minority

Note: * p < .05.
** p < .01.
Correlations for SAT include only those states (N=23) with at least 50% participation rate.
Correlations for ACT include only those states (N=25) with at least 50% participation rate.

**T**  **F**  **I**  1. Low SES appears to cause reduced achievement.

**T**  **F**  **I**  2. Income accounts for about 36% of the variance in ACT performance.

**T**  **F**  **I**  3. The relationship between SAT & Income is greater than the relationship between Grad. Rate % reduced lunches.

**T**  **F**  **I**  4. As class size goes down standardized test performance goes up.
Interpretive Exercises

The questions or problems which accompany the display may be presented in one or more of the following formats:

- Short answer or completion items
- Alternative response questions
- Matching exercises
- Multiple-choice questions
- Essay questions
Interpretive Exercises

Advantages:

• Allows you to test knowledge through the use of media that might be encountered in everyday situations (e.g., charts, pictures, etc.)
• Measure more complex knowledge through the analysis of data
• Greater depth
Interpretive Exercises

Disadvantages:

• Difficult to construct
• Must match test items with instructional goals and objectives
• Possibly a heavy demand on reading skill
Short answer/completion Questions

- **Advantages**
  - Measure **RECALL** instead of **RECOGNITION**—reduces guessing
  - Easy to construct

- **Disadvantages**
  - Not suitable for complex learning outcomes
  - More difficult scoring (spelling/handwriting/diversity of answers—e.g., Where was George Washington born?)
Writing GOOD short answer items:

- Generally, one blank per question
- Stem needs to indicate the expected response
- NOT:
  - IQ is _______
  - ___ is a better _____ of _____ than ______
- But:
  - The term IQ is an acronym for what phrase?
  - Regression is a better way of analyzing continuous data than ________

- Ask direct questions rather than incomplete statements--they are more natural to students and generally better structured
Essay items

The following are some objectives for which essay items are appropriate.

1. Draw warranted conclusions from evidence.
2. Interpret a written selection, art work, etc.
3. Decide what information is needed to solve a problem.
4. Formulate hypotheses.
5. Determine an author's frame of reference.
6. Express a reasoned judgment and defend it.
7. Compare and contrast methods, selections, authors, etc.
8. Identify the assumptions implicit in ...
9. Integrate knowledge from related areas.
10. Evaluate the worth, merit, utility of ...
Restricted vs. Extended Response

- Restricted -- limits both the content and response
  - Similar to what can be measured with an interpretive exercise but requires the supply rather than selection of the answer

- Extended -- allows the student to select, organize, integrate, and evaluate information
Linn & Miller on extended-response questions:

“Neither a hard-line measurement position that rejects extended essays as an approach to measurement nor one that ignores the difficulties of scoring seems to contribute much to the valid measurement of student achievement. It seems more sensible to identify the complex skills we want to measure, formulate questions that elicit these skills, evaluate the results as reliably as we can, and then use these admittedly limited data as the best evidence we have available.” (pg. 231)
Pros & Cons of Essays

- Measures learning outcomes that cannot be measured by other means
- Emphasis on higher level thinking skills
- Ease of construction ??
- Directly measures writing skills

- Difficult to score reliably
- Time consuming to score
- Limited sampling
Essay items

- Specify parameters and limitations:
  - In 100 words, one page, 3 paragraphs, etc.
  - List the advantages and disadvantages of...
  - Describe three reasons why the Nebraska Cornhuskers are the greatest college football team of all time.

- Try for several short essays rather than one large one – allows you to sample mastery of several objectives
- Be as specific in stem as possible
- If assessing mastery, be careful allowing choice of essays from list
- Don’t use essays to measure objectives more easily measured by objective items
Scoring essays

- If it is not in the objectives, don’t score it
  - Consider this statement when examining spelling/grammar
  - Handwriting/visual appeal should not count unless it is a design course
- Score anonymously if possible
- Score all essay #1 and then all essay #2, etc. for consistency
Develop a specific rubric to help you in awarding points:

- 5 points for successfully defining educational psychology
- 15 points for describing an application of operant conditioning successfully
- 15 points for evaluating a classroom in terms of Bloom’s Learning for Mastery
- 65 points for listing 5 reasons why Seinfeld was the greatest show ever created

Share rubric with STUDENTS before the test!
### SAT

<table>
<thead>
<tr>
<th></th>
<th>Old SAT</th>
<th>New SAT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>3 hours</strong></td>
<td></td>
<td><strong>3 hours 45 minutes</strong></td>
</tr>
<tr>
<td><strong>Critical Reading</strong></td>
<td><strong>75 minutes</strong></td>
<td><strong>70 minutes</strong></td>
</tr>
<tr>
<td></td>
<td>Two 30-minute sections and one 15-minute section</td>
<td>Two 25-minute sections and one 20-minute section</td>
</tr>
<tr>
<td><strong>Math</strong></td>
<td><strong>75 minutes</strong></td>
<td><strong>70 minutes</strong></td>
</tr>
<tr>
<td></td>
<td>Two 30-minute sections and one 15-minute section</td>
<td>Two 25-minute sections and one 20-minute section</td>
</tr>
<tr>
<td><strong>Writing</strong></td>
<td></td>
<td><strong>60 minutes</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Two multiple-choice sections (one 25-minute section and one 10-minute section) and one 25-minute essay</td>
</tr>
<tr>
<td><strong>Variable Section</strong></td>
<td><strong>30 minutes</strong></td>
<td><strong>25 minutes</strong></td>
</tr>
<tr>
<td></td>
<td>Old SAT</td>
<td>New SAT</td>
</tr>
<tr>
<td>----------------</td>
<td>-------------</td>
<td>-------------</td>
</tr>
<tr>
<td><strong>Critical Reading</strong></td>
<td>V 200–800</td>
<td>CR 200–800</td>
</tr>
<tr>
<td><strong>Math</strong></td>
<td>M 200–800</td>
<td>M 200–800</td>
</tr>
<tr>
<td><strong>Writing</strong></td>
<td></td>
<td>W 200–800</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Will include 2 subscores:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Essay 2–12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1/3 of writing score)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Multiple-choice 20–80</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2/3 of writing score)</td>
</tr>
</tbody>
</table>
New SAT Writing Section
Additional measure of an important college success skill Encourages writing in schools

Essay

- Students will read a short excerpt, or two quotations, and respond to a prompt that frames an issue.
- Students must first think critically about the issue presented in the essay assignment and then define and support their point of view, using reasoning and evidence based on their own experiences, readings, or observations.
- The essay will be similar to the type of on-demand writing that is typically done in college.
New SAT Writing Section

Prompts

- Prompts will be written to be easily accessible to the general test-taking population, including students for whom English is a second language (ESL), and to be free of figurative, technical, or specific literary references.

- Prompts will be relevant to a wide range of fields and interests, not narrowly related to specific topics.

- Prompts will be tested to ensure that they do not carry any bias across subgroups.
Essay Prompt

Think carefully about the issue presented in the following quotations and the assignment below.

1. While secrecy can be destructive, some of it is indispensable in human lives. Some control over secrecy and openness is needed in order to protect identity. Such control may be needed to guard privacy, intimacy, and friendship.

   Adapted from Sissela Bok, “The Need for Secrecy”

2. Secrecy and a free, democratic government, President Harry Truman once said, don’t mix. An open exchange of information is vital to the kind of informed citizenry essential to healthy democracy.

   Editorial, “Overzealous Secrecy Threatens Democracy”

Assignment: Do people need to keep secrets, or is secrecy harmful?

Plan and write an essay in which you develop your point of view on this issue. Support your position with reasoning and examples taken from your reading, studies, experience, or observations.
Essay Prompt

- Essays not written on the assigned topic will receive a subscore of zero for the essay portion of the writing section.
How Will the Essays Be Scored?

Readers will

- understand that the essay is a first draft;
- read quickly to gain an impression of the whole essay relative to the holistic Scoring Guide and the sample range-finder essays;
- read the entire essay before scoring and then score immediately;
- read supportively, looking for and rewarding what is done well rather than what is done badly or omitted;
- not judge an essay by its length or the quality of handwriting;
- understand that grammar is not an overriding factor in determining an essay score; and
- consider spelling only when errors are so persistent that they interfere with meaning.
How Will the Essays Be Scored?

- Readers for the new SAT writing section will be trained to recognize and reward a wide variety of writing styles and strategies for developing a point of view at each score point.
- The new SAT essay will neither reward nor punish formulaic approaches to writing, such as the five-paragraph essay.
- Prompts and the Scoring Guide call for directly relevant responses that cannot be coached or memorized ahead of test time.
## New SAT Scoring Guide

<table>
<thead>
<tr>
<th>Score of 6</th>
<th>Score of 5</th>
<th>Score of 4</th>
</tr>
</thead>
</table>
| An essay in this category is **outstanding**, demonstrating **clear and consistent mastery**, although it may have a few minor errors. A typical essay  
- effectively and insightfully develops a point of view on the topic and demonstrates outstanding critical thinking, using clearly appropriate examples, reasons, and other evidence to support its position  
- is well organized and clearly focused, demonstrating clear coherence and smooth progression of ideas  
- exhibits skillful use of language, using a varied, accurate, and apt vocabulary  
- demonstrates meaningful variety in sentence structure  
- is free of most errors in grammar, usage, and mechanics | An essay in this category is **effective**, demonstrating **reasonably consistent mastery**, although it will have occasional errors or lapses in quality. A typical essay  
- effectively develops a point of view on the topic and demonstrates strong critical thinking, generally using appropriate examples, reasons, and other evidence to support its position  
- is well organized and focused, demonstrating coherence and progression of ideas  
- exhibits facility in the use of language, using generally appropriate vocabulary  
- demonstrates some variety in sentence structure  
- is generally free of most errors in grammar, usage, and mechanics | An essay in this category is **competent**, demonstrating **adequate mastery**, although it will have lapses in quality. A typical essay  
- develops a point of view on the topic and demonstrates competent critical thinking, using adequate examples, reasons, and other evidence to support its position  
- is generally organized and focused, demonstrating some coherence and progression of ideas  
- exhibits adequate but inconsistent facility in the use of language, using generally appropriate vocabulary  
- demonstrates some variety in sentence structure  
- has some errors in grammar, usage, and mechanics |
## New SAT Scoring Guide

### Score of 3
An essay in this category is *inadequate*, but demonstrates *developing mastery*, and is marked by ONE OR MORE of the following weaknesses:

- develops a point of view on the issue, demonstrating some critical thinking, but may do so inconsistently or use inadequate examples, reasons, or other evidence to support its position
- is limited in its organization or focus, or may demonstrate some lapses in coherence or progression of ideas
- displays developing facility in the use of language, but sometimes uses weak vocabulary or inappropriate word choice
- lacks variety or demonstrates problems in sentence structure
- contains an accumulation of errors in grammar, usage, and mechanics

### Score of 2
An essay in this category is *seriously limited*, demonstrating *little mastery*, and is flawed by ONE OR MORE of the following weaknesses:

- develops a point of view on the issue that is vague or seriously limited, demonstrating weak critical thinking, providing inappropriate or insufficient examples, reasons, or other evidence to support its position
- is poorly organized and/or focused, or demonstrates serious problems with coherence or progression of ideas
- displays very little facility in the use of language, using very limited vocabulary or incorrect word choice
- demonstrates frequent problems in sentence structure
- contains errors in grammar, usage, and mechanics so serious that meaning is somewhat obscured

### Score of 1
An essay in this category is *fundamentally lacking*, demonstrating *very little or no mastery*, and is severely flawed by ONE OR MORE of the following weaknesses:

- develops no viable point of view on the topic, or provides little or no evidence to support its position
- is disorganized or unfocused, resulting in a disjointed or incoherent essay
- displays fundamental errors in vocabulary
- demonstrates severe flaws in sentence structure
- contains pervasive errors in grammar, usage, or mechanics that persistently interfere with meaning

---

*Essays not written on the essay assignment will receive a score of zero.*
Scoring Procedures for the Essay

- Procedures will be similar to those for the current SAT Subject Test in Writing.
- Essays will be scored by trained high school English teachers and college professors with experience teaching writing.
- Each essay will be scored independently by two readers according to the holistic Scoring Guide in conjunction with sample essays selected for training.
- Essays will be scored on a scale of 1 to 6 by each reader (total score of 2 to 12).
- Essays will be scanned and distributed to readers via the Web.
- Scoring and reader supervision will take place online.
Essays Will Be Scored Fairly and Accurately

- If the two readers’ scores differ by more than one point, the essay will be read by a third reader.
- Based on the College Board’s experience in scoring the SAT Subject Test in Writing, the rigorous reader training and qualification process, and continuous monitoring of readers as they score, the College Board expects that less than 8 percent of all essays will call for a third reader.
Performance Assessment

- Also referred to as authentic assessment or alternative assessment
- Process may be ill-defined in that there may be multiple effective solutions
- Best suited for the creation of a product (e.g., model city, painting, invention)
Performance Assessment discussion:

- Group grades?
- Adapt each exercise for student strengths?
- Easily seen and judged by others?
- Students creating the rubric?
Advantages of Performance Assessment

- Motivation increases
- Teach to “real life”
- Assessment and educational product are the same
- External validity
- Utilize energy
- Student understanding of their work increases
- Tap higher level thinking & problem solving
Disadvantages of Performance Assessment

- Often contrived assessments
- Increased cost, material
- Increased time
- More difficult to score reliably
- Capture only a small part of academic achievement
- More things to consider for classroom management and planning
Performance Assessment

Developing a good **rubric** is the key to having a quality performance assessment!
Scoring Rubrics

- Share with students beforehand
- Break them into meaningful facets/subscales related to learning objectives
- Weight each facet/subscale according to instructional emphasis & desired outcomes
- Clearly define score points throughout the scale
- Be able to clearly describe projects that vary from poor to excellent
Scoring Rubrics

Some common errors in scoring:

- **Bias errors** -- scoring is too generous, severe, or middle of the road
- **Halo effect** -- scoring influenced by knowledge of the student
- **Logical error** -- scoring influenced by beliefs about particular traits or characteristics (e.g., impulsive students more error prone)