“Look,” he said, “I found a limestone rock. I know it is a limestone rock because I found a rock last year that has the same color and it was limestone.”
Jesse said, “Just because it looks the same it doesn’t have to be the same.”

1. Which of the following explanations best supports Jesse’s point of view?
   A During the year the chemical properties of limestone probably changed.
   B Different minerals have very similar physical properties.
   C One year is not long enough for the minerals in a rock to change their physical properties.

Use the following information to answer Question 2.
Billy took the rock home and did an experiment with it. He put a piece of the rock in a clear glass and poured vinegar over it. The piece of rock bubbled and foamed. “There!” he said to Jesse, “That proves the rock is limestone.”
Jesse said, “No! You are wrong. You haven’t proved it!”

2. Why was Jesse correct?
   A Billy did the experiment only once. He needs to repeat the same type of experiment many times with different bits of the rock. If the mixture bubbles every time, that will prove it.
   B The experiment is correct but Billy misinterpreted the results. Limestone does not bubble and foam in vinegar.
   C Billy should do many different kinds of experiments, not just vinegar tests, because many different kinds of substances bubble and foam in vinegar.
   D Billy should not have used vinegar. He should have used distilled water. If the rock made the water warm, that would prove it is limestone.

Use the following information when answering Questions 3 to 6.
For a social studies project, a class interviewed all the 10th-grade students. They asked how many hours per week students worked at after-school jobs. They also asked what their average grades were last term. They found that students with Fs and Ds worked 8 to 10 hours per week, students with Cs and Bs worked 10 to 20 hours per week, and students with As worked 8 to 10 hours per week.

Alternative Format A
Students choose from among teacher-provided interpretations but are required to write a justification of their choice.

3. Which of the following is the most valid interpretation of these findings?
   A If you work 10 to 20 hours per week you will only get Cs and Bs.
   B Working after school is not related to your grades.
   *C A student who works 10 to 20 hours per week is probably not an A student.
   D The more hours a student works after school, the higher will be that student’s grades.

4. Write a brief explanation of why your answer to Question 3 is the most valid interpretation of these findings.

______________________________________________
______________________________________________
______________________________________________
______________________________________________

Alternative Format B
Students supply their own interpretation and justify it in writing.

5. What is the most valid interpretation of the relationship the class found between the number of hours students worked and their grades?

______________________________________________
______________________________________________
______________________________________________
______________________________________________

6. Write a brief explanation of why your interpretation of these findings is the most valid one.

______________________________________________
______________________________________________
______________________________________________
______________________________________________

The three variations (multiple-choice only, multiple-choice with short-answer, and short-answer only) in the social studies/mathematics example assess somewhat different abilities. Using multiple-choice only (Item 3) assesses a student’s ability to evaluate each of the interpretations you provide and select the best one. Thus you do not know a student’s reasoning behind his selection. The multiple-choice with short-answer combination (Items 3 and 4) assesses a student’s ability to explain or justify her choice from among the interpretations you provide as options. This helps you assess the reasoning behind students’ choices. The short-answer without the multiple-choice items (Items 5 and 6) assess both a student’s ability to interpret the experiment’s results and his ability to explain his reasoning. In this latter format, there may be multiple correct responses to the constructed-response questions. As with other
constructed-response items, you may want to give students partial credit if their response is not completely correct.

First, identify the learning targets you want to assess. The experiment-interpretation assessment format is appropriate when a learning target requires students to understand and interpret the results of empirical research. Before using this format for summative student evaluation, be sure you have taught and have given practice in interpreting the findings from empirical research studies.

Write the item to assess the student’s ability to apply specific principles. This means that you first identify the principles or rules you want students to apply, then craft the item so it requires students to use the principle in a new situation. For example, items in the preceding examples are crafted around the following principles:

- Different substances may share the same or similar physical properties such as color, texture, and solubility. [Item 1]
- Different substances may share the same or similar chemical properties, such as their reactivity with acids. [Item 2]
- Some patterns of relationships among variables are not strictly increasing or decreasing but are curvilinear. [Items 3 through 6]

After identifying the principle(s), you create the item in such a way that it requires students to use or apply the principle(s). Usually, this means writing a description of the experiment or research study that results in findings that a student can then interpret using the principle(s). (See the interpretive text that immediately precedes Items 1, 2, 3/4, and 5/6 in the previous examples.)

Next, draft a stem that asks the student to interpret or explain the experimental findings you describe. You may then list several correct or partially correct interpretations. You may also list incorrect interpretations that result from incomplete or faulty reasoning. Avoid using as distractors interpretations that are completely unrelated to the experiment you describe in the interpretive material or distractors that are “silly” or “tricky.” For example, it would be inappropriate for you to use in Item 1 a distractor such as “Jesse knows that Billy is a liar.”

As with the best-answer item format, distractors for this format should contain interpretations or explanations that contain your students’ typical misconceptions. To determine these misinterpretations, you could assign several open-ended questions as homework and select from among the students’ responses those that are excellent, good, and poor. Use these selections as a basis for creating multiple-choice options.

If you use the multiple-choice versions of this format, you should follow the basic rules of writing multiple-choice items that we discussed earlier and that are summarized in the multiple-choice checklist. If you use one of the short-answer versions of this format, you should follow the basic rules of short-answer item writing (Chapter 8). The following checklist offers specific guidance for the experiment-interpretation item format. Use it to review the items you craft.

A Checklist for Reviewing the Quality of Experiment-Interpretation Items

Ask these questions of every item you write. If you answer no to one or more questions, revise the item accordingly.

1. Does each item assess an important aspect of the unit’s instructional targets?
2. Does each experiment-interpretation item match your assessment plan in terms of performance, emphasis, and number of points?
3. Does each item focus on requiring students to apply one or more important principles or criteria to new situations, examples, or events?
4. Have you given students opportunity to practice applying the appropriate criteria or principles for judging the “best” or “most valid” interpretation?
5. Did you describe an experiment or research study in concise but sufficient detail that a student can use the appropriate criteria or principles to interpret the results?
6. Is the keyed answer the only one that can be defended as the “best” or “most valid” interpretation?
7. Is each distractor based on an important misconception, misinterpretation, or misapplication of a criterion or principle? Did you avoid tricky or trivial ways of making a distractor partially correct or contain misinformation?
8. Did you avoid (a) having more than one “best” or “most valid” answer and (b) using “all of the above” or “none of the above”?
9. Did you apply all of the appropriate item-writing guidelines described in the multiple-choice checklist?
Advantages  You may use the experiment-interpretation format to assess a student’s ability to evaluate explanations, interpretations, and inferences from data. The multiple-choice-only version allows you to score the items more quickly and more objectively than the other versions. Because students are required only to select the correct answer, their response times are shorter. Therefore, you can use more items and cover more content within a shorter assessment period than with short-answer items.

If the experiments and findings you present in the items are new to the students, your items will assess students’ ability to apply principles and criteria from your subject area. Using experiments and data new to your students in assessment tasks requires you to teach students how to apply criteria and principles to a variety of situations. You will need to give students sufficient practice in applying criteria and principles before assessing them for summative evaluation purposes. This will move your teaching away from teaching facts and results, and toward teaching students to actively apply their knowledge and skill.

If you require students to justify their multiple-choice answers, you will have some information about their reasoning processes. Students often make the correct choice from among the possible interpretations you give them, but they cannot explain why they made the choice, or they give faulty explanations. If you require a student both to supply his interpretation and to justify it, you can assess whether the student can generate and explain his own interpretations of experimental findings.

Criticisms  Like the best-answer item format, the experiment-interpretation format is not easy to write. You must know your subject matter and your students’ thinking patterns well enough to create items that allow you to identify faulty thinking as well as correct answers. Faulty thinking must be reflected in your multiple-choice distractors. This means you must be able to create partially correct interpretations and incorrect interpretations that people typically make.

Use experiment-interpretation items to assess higher-order thinking. Do not use this format to assess whether a student can remember the “correct” interpretations of specific experimental results you taught. Using this format to assess remembering encourages students to look to the teacher or the text as the source of fixed knowledge. It discourages students from learning skills required to interpret the empirical results of experiments.

Statement-and-Comment Items

A statement-and-comment item presents a statement about some relevant subject matter and requires the student either to write a comment about the statement or to select the most appropriate comment from among a list you provide. Here is an example of a statement-and-comment item:

Example

A. Multiple-choice version of a statement-and-comment item

The Bundle of Sticks—Aesop  An old man on the point of death summoned his sons around him to give them some parting advice. He ordered his servants to bring in a bundle of sticks, and said to his eldest son: “Break it.” The son strained and strained, but with all his efforts was unable to break the bundle. The other sons also tried, but none of them was successful. “Untie the bundle,” said the father, “and each of you take a stick.” When they had done so, he called out to them: “Now, break,” and each stick was easily broken. “You see my meaning,” said their father.

Directions: The quote expresses the theme of Aesop’s fable “The Bundle of Sticks.” Choose the answer that best expresses how the theme applies to the fable.

1. “Union gives strength.”
   A The three sons all tried to break the bundle.
   *B None of the sons could break the bundle of sticks.
   C Each of the sons could break a single stick.

B. Short-answer version of a statement-and-comment item

Directions: The quote expresses the theme of Aesop’s fable “The Bundle of Sticks.” Below the quote, explain how the theme applies to the fable.

2. “Union gives strength.”

   __________________________________________________________
   __________________________________________________________
In the multiple-choice version, a student selects from among several alternate choices the best meaning of the quoted theme. The multiple-choice version is a special case of the best-answer item format. The alternatives should be phrased in language different from the “pat phrases” learned in class. In the short-answer version, students must comment directly, writing their own interpretation of the quoted statement.

First, as always, identify the learning targets you want to assess. This assessment format is appropriate when a learning target requires a student to comprehend statements and themes.

If you give students the short-answer version as a homework exercise, you may use excellent, good, and poor student responses as a basis for creating the alternatives for the multiple-choice version. As with the best-answer variety, of which this may be considered a special case, you usually cannot use students’ responses verbatim as multiple-choice options; paraphrase them. Because the multiple-choice version of the statement-and-comment is a type of best-answer item, follow the guidelines suggested in the best-answer item checklist.

Advantages The statement-and-comment item format assesses a student’s ability to evaluate interpretations of a given statement. The multiple-choice version assesses whether students can identify the best interpretation from among several. Interpretations should not use the same wording used in class. Rather, they should be comments typically made by students when interpreting the quoted statement. In this way, students must rely on their comprehension of the quoted phrase instead of their memory of a “set” comment.

The open-ended version assesses students’ ability to recall and write about the meaning of the quoted statement. Although it may be an advantage to have students construct their own comments about the quoted statement, there is a downside. Students may just write an explanation or commentary they memorized from the class discussion or from a textbook. You have some control over what kinds of comments they must evaluate if you present the multiple-choice version.

Criticisms The statement-and-comment item format has limited applications. You must identify appropriate statements that students should interpret. Although there are many subjects for which such statements exist, the task itself represents a small range of learning targets. The short-answer version of the task does provide an opportunity for students to display their comprehension of the quoted statement. However, students may simply repeat the phrases they learned in class.

MATCHING EXERCISE FORMAT

A matching exercise presents a student with three things: (1) directions for matching, (2) a premise list, and (3) a response list. The student’s task is to match each premise with one of the responses, using as a basis for matching the criteria described in the directions. Figure 9.8 shows a matching exercise with its various parts labeled.

---

**Figure 9.8 Example of a matching exercise.**

**Directions:** In the left column below are descriptions of some late-19th-century American painters. For each description, choose the name of the person being described from the right column, and place the letter identifying it on the line preceding the number of the description. Each name in the right column may be used once, more than once, or not at all.

**Item numbers**

1. A society portraitist, who emphasized depicting a subject’s social position rather than a clear-cut characterization of the subject.
   - (d) 1. a. Mary Cassatt
   - b. Thomas Eakins
   - c. John LaFarge

2. A realistic painter of nature, especially known for paintings of the sea.
   - (d) 2. d. Winslow Homer
   - e. John Singer Sargent
   - f. James A. M. Whistler

3. A realistic painter of people, who depicted strong characterizations and powerful, unposed forms of the subject.
   - (b) 3. a. Mary Cassatt
   - b. Thomas Eakins
   - c. John LaFarge

4. An impressionist in the style of Degas, who often painted mother and child themes.
   - (a) 4. a. Mary Cassatt
   - b. Thomas Eakins
   - c. John LaFarge
The sample exercise in Figure 9.8 requires simple matching based on associations that students must remember. You may create matching exercises, however, to assess students’ comprehension of concepts and principles. Examples of these latter types appear later in the chapter.

In matching exercises, premises are listed in the left column and responses in the right column, or responses are listed vertically above the premises. Each premise is numbered because each is a separately scoreable item. Matching exercises can have more responses than premises, more premises than responses, or an equal number of each. Perfect matching occurs when you have an equal number of premise statements and response statements. Most assessment specialists consider perfect matching to be undesirable because, if a student knows four of the five answers, the student automatically gets the fifth (last) choice correct, whether or not he knows the answer. This reduces the validity of the assessment results.

Matching exercises are very much like multiple-choice items. Each premise functions as a separate item. The elements in the list of responses function as alternatives. You could rewrite a matching exercise as a series of multiple-choice items: Each premise would then be a multiple-choice stem, but the same alternatives would be repeated for each of these stems. This leads to an important principle for crafting matching exercises: Use matching exercises only when you have several multiple-choice items that require repeating the identical set of alternatives.

ADVANTAGES AND CRITICISMS OF MATCHING EXERCISES

Advantages

A matching exercise can be a space-saving and objective way to assess a number of important learning targets, such as your students’ ability to identify associations or relationships between two sets of things. You can also develop matching exercises using pictorial materials to assess the students’ abilities to match words and phrases with pictures of objects or with locations on maps and diagrams. Figure 9.9 gives examples of relationships that you may use as a basis for developing matching exercises.

Criticisms

Detractors criticize the matching exercise because students can use rote memorization to learn the elements in two lists, and because teachers often use matching exercises only to assess such rote associations as names and dates. As a result, critics often see this assessment format as limited to the assessment of memorized factual information.

Thoughtful teachers, however, also use matching exercises to assess aspects of students’ comprehension of concepts, principles, or schemes for classifying objects, ideas, or events (we will see examples later). If you want to assess students on these higher-level abilities, create exercises that present new examples or instances of the concept or principle to the students. Then require students to match these examples with the names of appropriate concepts or principles. In this context, new examples are instances of concepts that students have not been previously taught or encountered. Similarly, a matching task can describe a situation novel to the student, and the student can decide which of several rules, principles, or classifications is likely to apply. An example of this type of matching exercise follows:

Example

Directions: Each numbered statement below describes a testing situation in which ONE decision is represented. On the blank next to each statement, write the letter:

A if the decision is primarily concerned with placement
B if the decision is primarily concerned with selection
C if the decision is primarily concerned with program improvement
D if the decision is primarily concerned with theory development
E if the decision is primarily concerned with motivating students

(A)1. After children are admitted to kindergarten, they are given a screening test to determine which children should be given special training in perceptual skills.

<table>
<thead>
<tr>
<th>Possible premise sets</th>
<th>Associated response sets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accomplishments</td>
<td>Persons</td>
</tr>
<tr>
<td>Noted events</td>
<td>Dates</td>
</tr>
<tr>
<td>Definitions</td>
<td>Terms and phrases</td>
</tr>
<tr>
<td>Examples, applications</td>
<td>Rules, principles, and classifications</td>
</tr>
<tr>
<td>Concepts (ideas, operations, quantities, and qualities)</td>
<td>Symbols and signs</td>
</tr>
<tr>
<td>Titles of works</td>
<td>Authors and artists</td>
</tr>
<tr>
<td>Foreign words and phrases</td>
<td>English correspondence</td>
</tr>
<tr>
<td>Uses and functions</td>
<td>Parts and machines</td>
</tr>
<tr>
<td>Names of objects</td>
<td>Pictures of objects</td>
</tr>
</tbody>
</table>