

CHAPTER

5

Selected Response Assessment

S elected response assessment has come to signify the traditional, old-fashioned way of testing. As recently as the 1920s and 1930s, however, it was known as the new, scientific assessment method, welcomed by those at the forefront of education innovation because it was considered objective—that is, free of teacher judgment. Although its luster has dimmed over the years, it is still a valuable tool in our assessment repertoire and worthy of regular use as both assessment *for* learning and assessment *of* learning.

Selected response assessments can include one or more of four different item types: multiple choice, true/false, matching and short answer fill in. In this chapter, we will examine the following:

- When to use selected response assessments.
- How to build high-quality tests.
- How to use them as assessments for learning by involving students in their planning, development, and use.

The shaded areas of Figures 5.1 and 5.2 show where we are within our framework of study.

Figure 5.1 Keys to Quality Classroom Assessment

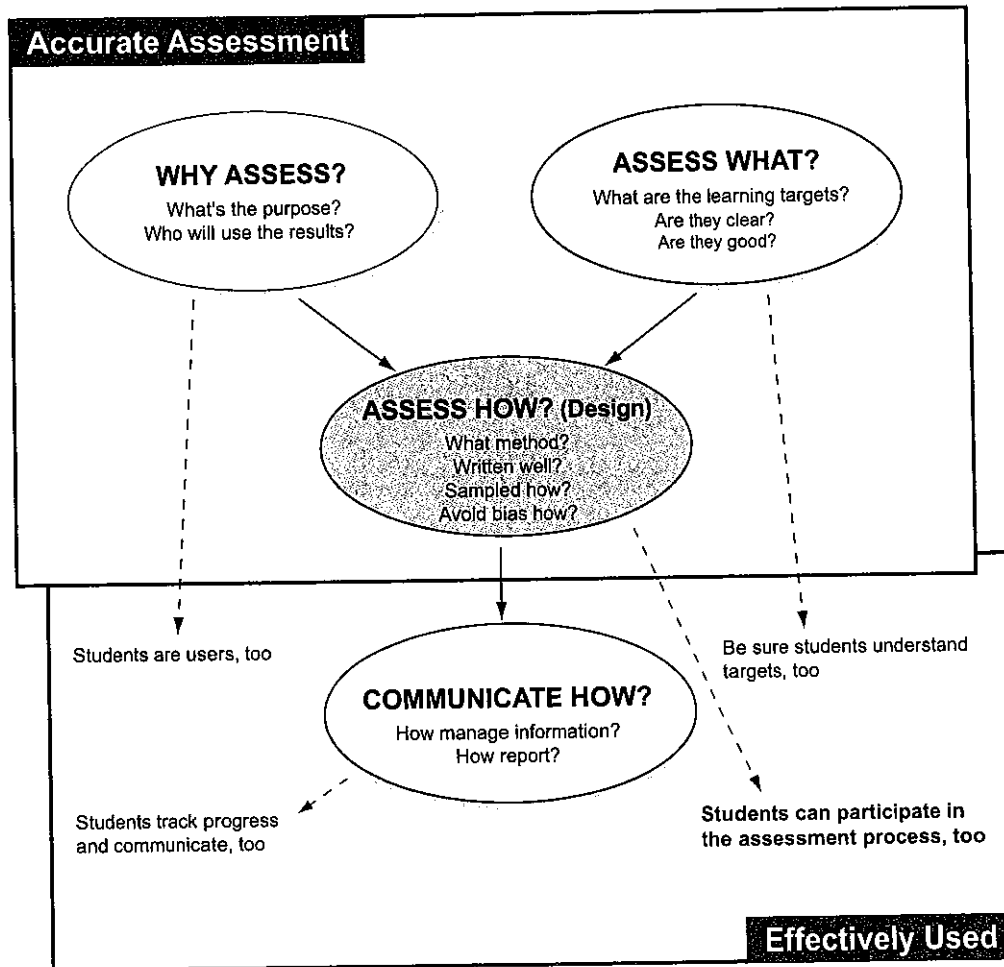


Figure 5.2 A Plan for Matching Assessment Methods with Achievement Targets

Target to Be Assessed	Assessment Method			
	Selected Response	Extended Written Response	Performance Assessment	Personal Communication
Knowledge Mastery				
Reasoning Proficiency				
Performance Skills				
Ability to Create Products				

Source: Adapted from *Student-Involved Assessment for Learning*, 4th ed. (p. 84), by R. J. Stiggins, 2005, Upper Saddle River, NJ: Merrill/Prentice Hall. Copyright © 2005 by Pearson Education, Inc. Reprinted by permission of Pearson Education, Inc.

When to Use Selected Response Assessment

Matching Method to Target

The first condition for using selected response is the type of learning target to be assessed, as described in Activity 4.4, “Target–Method Match,” in Chapter 4. Selected response formats are ideal for assessing knowledge-level learning targets and some patterns of reasoning, as shown in Figure 5.2.

TRY THIS

Activity 5.1 Learning Targets Best Assessed with Selected Response

To begin applying the content of this chapter to your own context, identify five to seven knowledge and/or reasoning learning targets in the curriculum materials you use. Write them down and save them for further activities.

Other Contextual Conditions

Several other key conditions influence choosing the selected response method of assessment:

- Students can read English well enough to understand what each test item is asking of them.
- The content to be assessed is broad, requiring wide-ranging coverage. Since the response time to one item is so short, we can include lots of items per unit of testing time and thus sample student achievement thoroughly.
- The assessment is to be taken by a large number of students. Electronic or optical scan scoring of response sheets increases scoring efficiency and accuracy.

Developing Selected Response Tests

As we saw in Chapter 4, each individual classroom assessment represents a part of a long-term assessment map that parallels the curriculum map for the reporting period. Each assessment contributes to an accumulating body of evidence of each student's level of achievement. Some assessments will support learning, others will serve to verify that learning has happened. Within this big picture, we plan each individual assessment, some of which may be selected response tests.

We will follow the five stages in the assessment development cycle described in Chapter 4 to develop selected response assessments:

1. Plan the assessment.
2. Develop the assessment.
3. Critique the assessment.
4. Administer the assessment.
5. Revise the assessment.

TRY THIS

Activity 5.2 Create a Quiz

Refer to the learning targets you listed in Activity 5.1. Identify those that you might assess together on one test, and then locate the written material—your textbook or other print material—you use to teach to those learning targets.

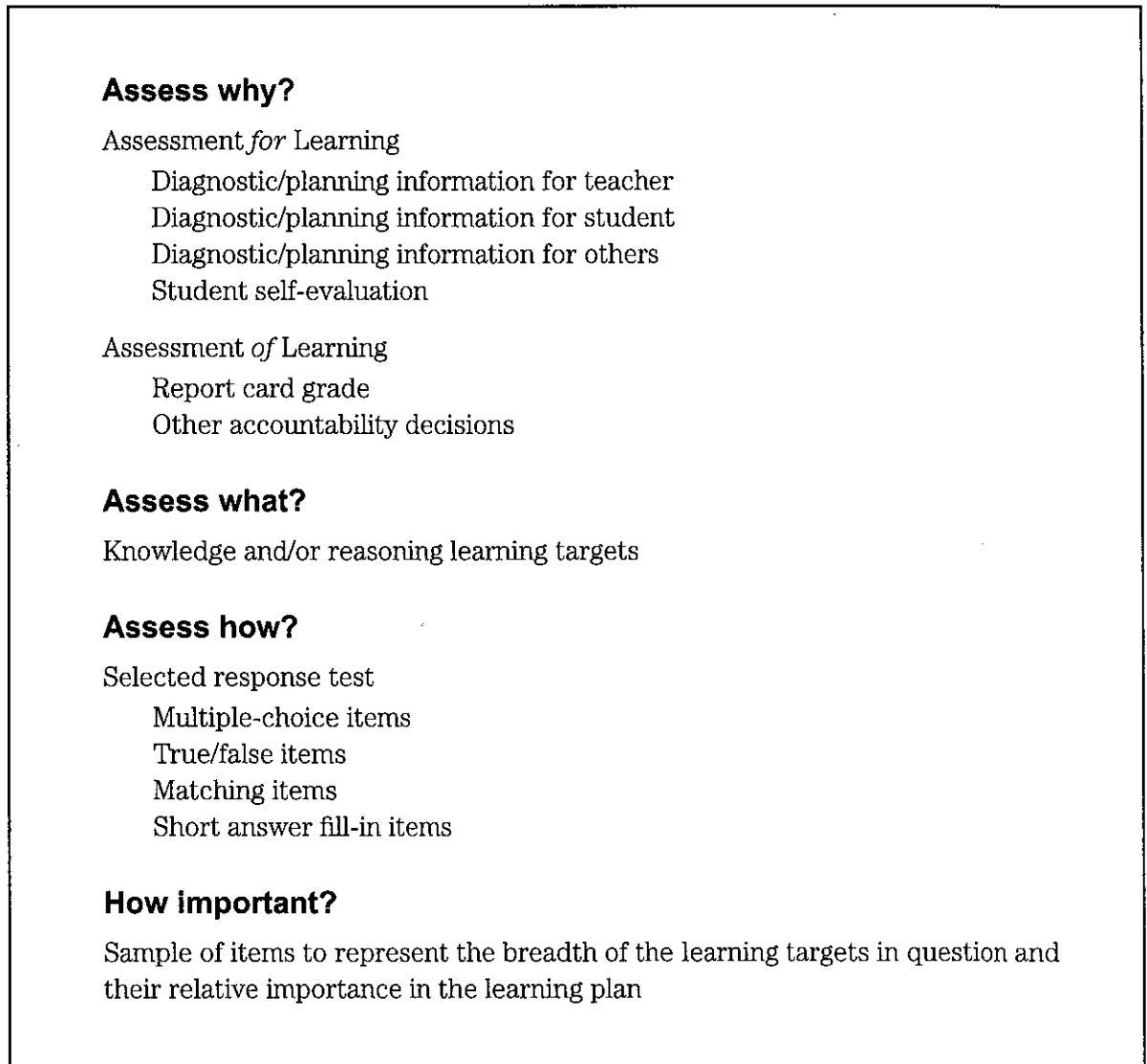
- Select a section of that material that is long enough (or extensive enough in its coverage) to permit you to build a 10-item selected response quiz.
- Review the material and build the quiz, being sure to include at least one instance of each selected response format (multiple choice, true/false, matching, and short answer fill in).

Set your new quiz aside for later use and continue reading.

Stage 1: Plan the Assessment

In the planning stage, as presented in Chapter 4, we answer four questions: Assess why? Assess what? Assess how? How important? These decisions are reflected in Figure 5.3. You may want to refer to the CD file, "Test Planning Forms," as you proceed through this first stage.

Figure 5.3 Decisions at the Planning Stage for a Selected Response Assessment



Assess Why?

This first question establishes the purpose for the assessment. It requires that we answer the following questions: How do we want to use the information? Who else will use it? What decisions will they make? For example, if we decide students are one of the intended users of the information, that will drive certain design decisions along the way, as you will see.

Assess What?

This question requires that we list the specific learning targets the test is to measure.

Assess How?

The answer here is, “by means of a selected response test,” so we must make sure we have identified only knowledge and reasoning learning targets as the subject of this test. (Although the kinds of selected response items available are listed in Figure 5.3, we need not choose which to use until we develop the test.)

How Important?

This question requires that we assign a relative importance to each learning target. One simple way to do this with selected response questions is to decide how many points the test will be worth and then divide the points according to relative importance of each learning target. The number of points we assign to each learning target outlines our sample. As described in Chapter 4, we select a sample of items to represent the breadth of the learning targets and their importance relative to each other in the instructional period the test is to cover.

Remember, when identifying the relative importance of each learning target, we consciously match our emphasis in assessment to our emphasis in the classroom. If, say, we spend 50 percent of the time learning how to read maps, then roughly 50 percent of the assessment should focus on map reading. If only 5 percent of the course deals with reading maps, then it would be inappropriate to spend 50 percent of the final assessment on map reading.

When using a test developed by someone else—another teacher, a textbook company, or a test publisher—it is crucial to examine it carefully and adjust it as needed for both the learning targets represented and the amount of emphasis each receives prior to deciding to use the test, using the procedure given in Activity 4.4, “Analyze Your Own Assessment for Clear Targets.”

Options for Test Plans

As you recall from Chapter 4, in the classroom we propose two useful types of test plans. One is a list of the learning targets and the other is a table representing the learning targets in shortened form. Both are equally effective as test planning instruments.

Table 5.1 shows a plan for a third-grade mathematics test consisting of a list of learning targets.

Table 5.1 Sample Test Plan for a Third-Grade Mathematics Test

Learning Targets	Importance
<i>Number Sense:</i> Identifies place value to thousands Reads, writes, orders, and compares numbers through four digits Reads and writes common fractions to represent models, real-life objects, and diagrams	11 points
<i>Number Operations:</i> Subtracts whole numbers to three digits with borrowing	4 points
<i>Measurement:</i> Reads time to the nearest minute Identifies correct units of measurement for length, capacity, weight, and temperature	5 points

There are times when we may want to frame our test plan in a table format, which is especially useful if we want to ensure that the test covers both recall of important information and reasoning processes we have taught. Figure 5.4 and Table 5.2 show a simple example of a list of learning targets and a test plan for a fifth-grade social studies unit on westward expansion. (Remember that there could be other learning targets taught during the unit—this test plan represents only those covered by the selected response portion of the test.)

Figure 5.4 Assess What? Fifth-Grade Social Studies Unit Learning Targets to Be Assessed by a Selected Response Test

- Explain the concept of Manifest Destiny and its contribution to the migration of people in the development of the United States.
- Compare the motives of the different groups who participated in the westward expansion by leaving the eastern United States and heading west.
- Compare the lives of different Native American groups before and after westward expansion.
- Identify significant individuals who took part in the westward expansion.
- Explain how the westward migration led to conflict between Native Americans and settlers and between Mexicans and settlers.

Table 5.2 *How Important?* Selected Response Test Plan for Fifth-Grade Social Studies Unit

Content Category	Know	Compare/ Contrast	Totals
Manifest Destiny	2		2
Reasons settlers went west	4	2	6
Life of Native American groups	4	2	6
Significant individuals	4		4
Conflicts caused by westward expansion	10		10
Effects on Native Americans and Mexicans	6		6
TOTALS	30	4	34

In the column labeled *Content Category* on the left of Table 5.2, the learning targets are translated into phrases, or categories, representing the content of each. Each category represents many facts and concepts, some of which will be sufficiently important to test. Table 5.2 also includes columns labeled for the cognitive action to be carried out: know outright and reason comparatively. These patterns will be emphasized during the unit of study. The numbers in each cell represent its relative importance in the unit as planned.

Stage 2: Develop the Assessment

So far, our test plan has yielded a list of learning targets or important concepts and an indication of their relative importance. The steps we take in this next stage are (1) determining specifically what to test, (2) determining what kinds of items to write and writing them, and (3) assembling the test.

TRY THIS**Activity 5.3 Make a Test Plan for Your Quiz**

Review the material out of which you built the 10-item quiz in Activity 5.2. If you were to build a test plan reflecting the priorities of that content, what would it look like? Please develop that plan now. You may want to use the form on the CD from Activity 4.4, “Analyze for Clear Targets.”

Identifying Important Elements

Even though we have numbers on the test plan to indicate the relative importance of each learning target or content category, we still need to identify what content we will test for each cell. In most cases, we won't test everything students are to have learned. Instead, we will select or create questions that cover as much of the important content as possible, given the amount of testing time available, and that are prudent for the age of our students. We use the results to make an inference: a student who has scored 75 percent on the test has mastered about 75 percent of the material that was intended to be learned. We must carefully select the subset of all possible important aspects of knowledge and reasoning so that our sample allows us to estimate level of achievement accurately.

Even though people often think of selected response tests as objective measures of learning, selecting the content for the test is itself a subjective exercise. The test developer—you yourself, a textbook author, or a test publisher—chooses what will be on the test from a vast array of possibilities. It is a matter of professional judgment, just as is determining how to teach the material in the first place. This element of subjectivity does not compromise the test's validity if we have clearly and accurately identified the learning targets that underpin our content standards.

If we are developing a test in a content area such as social studies or science, we first identify the specific content to include on the test by writing propositions. Propositions are statements of important facts, concepts, or understandings that we will be teaching students to know and understand. They state important elements of content and the kind of reasoning to be carried out, and function as basic units of any kind of selected response item we wish to write.

To write propositions, we begin by reviewing the material we have taught. For every cell in our test plan, we note in writing the most important facts, concepts, or understandings we think every student should have at the end of instruction. We use clearly worded sentences, and write down more propositions than we will need. Additional propositions serve two purposes: (1) they allow the writer to create parallel forms of the test; and (2) because some propositions may not lend themselves to writing clearly focused selected response items, we can decide to assess them with a different method, if needed.

Knowledge Propositions

If, for example, we are writing propositions for the test planned in Table 5.2, we will need a total of 30 knowledge items, 2 of which will relate to Manifest Destiny. As we read through the material, we identify and write down three or four statements that reflect important knowledge about the concept of Manifest Destiny. These are our propositions. They might include the following:

- Manifest Destiny represents a belief that it was natural and right to expand the territory of the United States westward.
- Manifest Destiny represents a mission to impart the government and way of life of United States citizens to people living in the land west of the United States during the 1800s.
- Manifest Destiny was first used as a justification for annexing Texas to the United States.
- Manifest Destiny represents a belief that was used to justify the taking of Native American lands.

The test plan also requires six items in the cell that crosses *Know* with *Effects on Native Americans and Mexicans*. Here are two sample propositions:

- Three effects of westward expansion on Plains Indians were increased disease, removal to reservation lands, and loss of food sources.
- Mexico lost the territory of Texas.

Reasoning Propositions

A reasoning proposition states the result of reasoning applied accurately to the information at hand. To write one, we identify the knowledge to be applied, apply the pattern of reasoning, and state the result as a declarative sentence. Propositions for the cell in Table 5.2 that crosses *Compare/Contrast* with *Reasons settlers went west* might read like this:

- The Mormons went west to practice their religion without persecution, whereas the settlers in Texas went west because land was cheap or free and they wanted a place to start over.
- Both Mormons and settlers in Texas were searching for a better life.
- Settlers were encouraged to move to Texas by the Mexican government, while Mormons were led to settle in Utah by their religious leaders.

Remember that when we intend to evaluate students' ability to reason, we must provide them with a context different than that in which they practiced. If we don't, as we saw in Chapter 3, we will not be capturing real evidence of their reasoning (ability to figure things out). Instead, we will have information about what they remember. If we want to assess the learning target, "Compare the motives of the different groups who participated in the westward expansion by leaving the eastern United States and heading west," we cannot have students practice comparing and contrasting during instruction using the same examples we will use on the test.

TRY THIS

Activity 5.4 Writing Propositions for Your Quiz

Refer to the 10-item test you created in Activity 5.2 and the material that you used as its basis. Identify the 10 propositions that underpin the items on your quiz. Do these adequately represent the most important learnings, in your opinion?

Writing Items

Once you have identified the propositions that reflect important learning, you can write any kind of selected response item you might want to use: multiple choice, true/false, matching, or short answer fill-in. Here's how it works with the following proposition from the Manifest Destiny example:

Manifest Destiny represents a mission to impart the government and way of life of United States' citizens to people living in the land west of the United States during the 1800s.

Multiple-Choice Items

To create a multiple-choice item, begin with a question. Then add a number of answers, only one of which is correct.

- What was the mission of Manifest Destiny in the United States in the 1800s?*
- a. To have Lewis and Clark make friends with the Native Americans they met.*
 - b. To move the U.S. form of government and way of life west.*
 - c. To defeat General Santa Anna in the battle of the Alamo.*
 - d. To establish religious freedom for all who lived in the west.*

Figure 5.5 illustrates another example of a proposition from the same social studies unit used to generate each of the different item types.

True/False Items

To create a true/false item that is true, include the proposition on the test as stated. (In this example, for fifth graders, you may want to simplify the proposition so that it reads as follows: Manifest Destiny represents a mission the U.S. had in the 1800s to move its government and way of life westward.)

To create a false true/false item, make one part false:

Manifest Destiny represents a mission the United States had in the 1800s to guarantee religious freedom to all settlers.

Figure 5.5 Turning a Proposition into Different Item Types

<p>Proposition: Three effects of westward expansion on Plains Indians in the 1800s were increased disease, removal to reservation lands, and loss of food sources.</p>
<p>True/False Item: (True) Three effects of westward expansion on Plains Indians were increased disease, removal to reservation lands, and loss of food sources. (False) One effect of westward expansion on Plains Indians was access to better health care.</p>
<p>Fill-in or Short Answer Item: What were three effects of westward expansion on Plains Indians in the 1800s?</p>
<p>Multiple-choice Item: What were three effects of westward expansion on Plains Indians in the 1800s? a. Access to health care, removal to reservation lands, and loss of food sources b. Access to health care, population growth, and opportunities for better jobs c. Increased disease, removal to reservation lands, and loss of food sources d. Loss of their schools, removal to reservation lands, and private ownership of land</p>

Matching Items

A matching exercise is similar to a multiple-choice item, in that the task is to combine the trigger item (or “stem”) with its proper match. To identify the content of the trigger item and the match, you simply take a proposition and separate it into to subject and predicate parts. The context where matching items makes sense is where the learning targets can be thought of as a series of closely linked propositions, such as states and their capitals or items to be categorized and their categories. Any individual match (stem and response) would state a single proposition. Matching items generally test knowledge propositions, but they can also be used to assess reasoning propositions.

Fill-in or Short Answer Items

To create a fill-in item, leave out the phrase defining the concept or dealing with the effect and ask a question:

What was the mission of Manifest Destiny in the United States in the 1800s?

Selecting from Among the Formats

Each of the item types has its proper uses. Table 5.3 details strengths and weaknesses of multiple-choice, true/false, matching, and short answer fill-in, test formats.

Table 5.3 Comparison of Selected Response Item Types

ITEM TYPE	USED WHEN	ADVANTAGE	LIMITATIONS
Multiple Choice	There is only one right answer. There are several plausible alternatives to the correct answer.	Can measure a variety of objectives. Easy to score. Can cover lots of material efficiently. Carefully crafted distracters can provide diagnostic information.	Guessing can skew score (up to 33% chance, depending on number of distracters). Can be hard to identify plausible distracters.
True/False	A large domain of content is to be tested, requiring the use of many test items.	Can ask many questions in a short time. Easy to score.	Can be trivial or misleading if not written carefully. Guessing can skew score (50% chance).
Matching	There are many related thoughts or facts; you want to measure association of information.	Can cover lots of material efficiently. Easy to score. Can serve as several multiple-choice items in one (each response is a distracter for the others).	Process of elimination can skew score if not written carefully.
Short Answer or Fill in the Blank	A clear, short answer is required. You want to determine if students know the answer, rather than if they can select it from a list.	Assesses production of a response. Reduces the possibility of getting the right answer by guessing. Can cover lots of material efficiently.	Takes longer to score.

Guidelines for Writing Quality Items

We offer here the commonsense guidelines that test developers use to ensure item quality.¹ The first set of guidelines applies to all item types, and the rest are specific to each particular format. Before reading them, you may want to take the short test in Activity 5.5.

TRY THIS

Activity 5.5 Franzipanics

Imagine you are a student in a class that has just studied the topic of “Franzipanics” and now it’s test time. Frankly, this is not your best subject, you don’t enjoy it, and you have not studied a lick. However, because you are pretty good at figuring things out, you are prepared to see what you can do on this test without knowing the content. Take the Franzipanics test in Figure 5.6 now. As you do, keep track of how you are figuring out the right answers. After taking the test, go to the file on the CD, “Franzipanics Answers,” and compare your answers to those in the answer key. Then read through the guidelines for test and item quality below to see which one or ones each item on the Franzipanics test addresses.

General Guidelines

1. *Keep wording simple and focused. Aim for the lowest possible reading level. Good item writing represents an exercise in effective written communication.*

Not this:

When scientists rely on magnets in the development of electric motors they need to know about poles, which are?

But this:

What are the poles of a magnet called?

- a. *Anode and cathode*
- b. *North and south*
- c. *Strong and weak*
- d. *Attract and repel*

Figure 5.6 Test of Franzipantics

Directions: Circle the correct answer for each question.

1. The purpose of the cluss in furnpaling is to remove
 - a. cluss-prags
 - b. tremalis
 - c. cloughs
 - d. plumots

2. Trassig is true when
 - a. lusp trasses the vom
 - b. the viskal flans, if the viskal is donwil or zortil
 - c. the belgo frulls
 - d. dissles lisk easily

3. The sigla frequently overfesks the trelsum because
 - a. all siglas are mellious
 - b. siglas are always votial
 - c. the trelsum is usually tarious
 - d. no trelsa are feskable

4. The fribbled breg will minter best with an
 - a. derst
 - b. morst
 - c. sorter
 - d. ignu

5. Among the reasons for trystal doss are
 - a. the sabs foped and the foths tinzed
 - b. the kredges roted with the orots
 - c. few rakobs were accepted in sluth
 - d. most of the polats were thonced

6. Which of the following (is, are) always present when trossels are being gruvен?
 - a. rint and vost
 - b. sot and plone
 - c. shum and vost
 - d. vost

7. The mintering function of the ignu is most effectively carried out in connection with
 - a. a raxma tol
 - b. the groshing stantol
 - c. the fribbled breg
 - d. a frally sush

8.
 - a.
 - b.
 - c.
 - d.

Source: From *Practice with Student-Involved Classroom Assessment* (p. 126), by J. A. Arter & K. U. Busick, 2001, Portland, OR: Assessment Training Institute. Copyright © 2006, 2001 Educational Testing Service. Reprinted by permission.

2. *Ask a full question in the stem.* This forces you to express a complete thought in the stem or trigger part of the question, which usually promotes students' understanding.

Not this:

Between 1950 and 1965

- a. *Interest rates increased.*
- b. *Interest rates decreased.*
- c. *Interest rates fluctuated greatly.*
- d. *Interest rates did not change.*

But this:

What was the trend in interest rates between 1950 and 1965?

- a. *Increased only*
- b. *Decreased only*
- c. *Increased, then decreased*
- d. *Remained unchanged*

3. *Eliminate clues to the correct answer either within the question or across questions within a test.* When grammatical clues within items or material presented in other items give away the correct answer, students get items right for the wrong reasons.

Not this:

All of these are an example of a bird that flies, except an

- a. *Ostrich*
- b. *Falcon*
- c. *Cormorant*
- d. *Robin*

(The article an at the end of the stem requires a response beginning with a vowel. As only one is offered, it must be correct.)

Not this either:

Which of the following are examples of birds that do not fly?

- a. Falcon*
- b. Ostrich and penguin*
- c. Cormorant*
- d. Robin*

(The question calls for a plural response. As only one is offered, it must be correct.)

4. *Do not make the correct answer obvious to students who have not studied the material.*
5. *Highlight critical, easily overlooked words (e.g., NOT, MOST, LEAST, EXCEPT).*
6. *Have a qualified colleague read your items to ensure their appropriateness. This is especially true of relatively more important tests, such as big unit tests and final exams.*
7. *Double check the scoring key for accuracy before scoring.*

Guidelines for Multiple-Choice Items

The following guidelines for writing multiple-choice test items allow students to answer questions more quickly without wasting time trying to determine what the question is saying. Here's some multiple-choice lingo. The item "stem" refers to the part of the question that comes before the choices. The "distracters" are the incorrect choices.

1. *Ask a complete question to get the item started, if you can.* This has the effect of placing the item's focus in the stem for clarity, not in the response options.
2. *Don't repeat the same words within each response option; rather, reword the item stem to remove the repetitive material from below.* This will clarify the problem and make it more efficient for respondents to read.
3. *Be sure there is only one correct or best answer.* This is where that colleague's independent review can help. Remember, it is acceptable to ask respondents to select a "best answer" from among a set of correct answers. Just be sure to word the question so as to make it clear that they are to find the best answer.

4. *Word response options as briefly as possible and be sure they are grammatically parallel.* This makes items easier to read and eliminates cues to the right answer.

Not this:

Why did colonists come to the United States?

- a. *To escape heavy taxation by their native governments*
- b. *Religion*
- c. *They sought the adventure of living among native Americans in the new land*
- d. *There was the promise of great wealth in the New World*
- e. *More than one of the above answers*

But this:

Why did colonists migrate to the United States?

- a. *To escape taxation*
- b. *For religious freedom*
- c. *For adventure*
- d. *More than one of the above*

5. *Make all response options the same length.* Testwise students know that the correct answer may be the longest one because writers frequently need to add qualifiers to make it the best choice. If you need to do this, do it to all response options.
6. *Don't use "all of the above" or "none of the above" merely to fill space; use them only when they fit comfortably into the context of the question. In general, test writers avoid using "all of the above" because if a student can determine that two responses are correct, then the answer must be "all of the above."*
7. *Use "always" or "never" in your answer choices with caution.* Rarely are things always or never true. Absolutes are frequently incorrect; a student who knows this but is not sure of the correct answer can automatically eliminate those choices.
8. *It's okay to vary the number of response options presented as appropriate to pose the problem you want your students to solve.* While four or five response options are most common, it is permissible to vary the number of response options offered across items within the same test. It is more important to have plausible distracters than a set number of them.

By the way, here's a simple, yet very effective, multiple-choice test item writing tip: If you compose a multiple-choice item and find that you cannot think of enough plausible distracters, include the item on a test the first time as a fill-in question. As your students respond, those who get it wrong will provide you with a good variety of viable distracters.

Guidelines for True/False Exercises

You have only one simple guideline to follow here: Make the item entirely true or false as stated. Complex "idea salads" including some truth and some falsehood just confuse the issue. Precisely what is the proposition you are testing? State it and move on to the next one.

Not this:

From the Continental Divide, located in the Appalachian Mountains, water flows into either the Pacific Ocean or the Mississippi River.

But this:

The Continental Divide is located in the Appalachian Mountains.

Guidelines for Matching Items

When developing matching exercises, follow all of the multiple-choice guidelines offered previously. In addition, observe the following guidelines:

1. *Provide clear directions for making the match.*
2. *Keep the list of things to be matched short.* The maximum number of options is 10. Shorter is better.
3. *Keep the list of things to be matched homogeneous.* Don't mix events with dates or with names.

Not this:

- | | |
|-----------------|----------------|
| ___ 1. Texas | A. \$7,200,000 |
| ___ 2. Hawaii | B. Chicago |
| ___ 3. New York | C. Mardi Gras |
| ___ 4. Illinois | D. Austin |
| ___ 5. Alaska | E. 50th state |

But this:

Directions: New England states are listed in the left-hand column and capital cities in the right-hand column. Place the letter for the capital city in the space next to the state in which it is located. Responses may be used only once.

<u>States</u>	<u>Capital Cities</u>
_____ 1. Rhode Island	A. Concord
_____ 2. Maine	B. Boston
_____ 3. Massachusetts	C. Providence
_____ 4. New Hampshire	D. Albany
_____ 5. Vermont	E. Augusta
	F. Montpelier

4. *Keep the list of response options brief in their wording and parallel in construction.*
5. *Include more response options than stems and permit students to use response options more than once when appropriate.* This has the effect of making it impossible for students to arrive at the correct response purely through a process of elimination.

Guidelines for Fill-in Items

Here are three simple guidelines to follow:

1. *Ask respondents a question and provide space for an answer.* This forces you to express a complete thought.
2. *Try to stick to one blank per item.* Come to the point. Ask one question, get one answer, and move on to the next question.

3. *Don't let the length of the line to be filled in be a clue as to the length or nature of the correct response.* This may seem elementary, but it happens. Again, this can misinform you about students' real levels of achievement.

Not this:

*In the percussion section of the orchestra are located _____,
_____, _____, and _____.*

But this:

In what section of the orchestra is the kettle drum found?

4. *Put the blank toward the end of the sentence.*

Interpretive Exercises

This is a label used for those instances when we present students with a table of information, a diagram, or some other source of information and then ask them to use that information to figure out answers to reasoning questions. The most common version of this is found in reading comprehension tests, where a passage is accompanied by test items that ask for inferences based on the content of the passage. You might consider using this exercise format when you are not sure that some (or all) of your students have mastered some body of basic knowledge, but nevertheless, you want to assess their reasoning proficiency. Or, you simply want to assess reasoning and don't need to assess content knowledge. Just give the content knowledge and ask them to use it. Table 5.4 shows examples of what such items might look like. For more examples, go to the file on the CD, "Reasoning Item Formulas."

Amount of Testing Time to Allow

Estimate time required to give the test as planned. The typical sitting should be no longer than 45 minutes in grades 3–12. Students can answer anywhere from 25 to 90 questions per sitting depending on their complexity and the students' grade level. Multiple-choice items typically take between 30 and 60 seconds to answer.

Table 5.4 Reasoning Test Item Formulas—Reading Comprehension

Reasoning Learning Target	Item Formula
Make inferences based on the reading	<p>Which sentence tells an idea you can get from this (selection)?</p> <p>a. The correct response is an idea that can reasonably be inferred from the text.</p> <p>b. Incorrect responses are ideas that it seems one might infer from the text but that the selection does not really support.</p>
Make predictions based on the reading	<p>What do you think (character) will do now that (cite circumstances at end of story)?</p> <p>a. The correct response is an outcome that can reasonably be predicted given the information in the text.</p> <p>b. Incorrect responses are not appropriate given the information in the text.</p>
Compare and contrast elements of text	<p>Which sentence tells how (two characters in the story) are alike?</p> <p>a. The correct response identifies an appropriate similarity.</p> <p>b. Incorrect responses do not identify similarities; they may focus on something that is true of one character or the other but not both.</p>
Make connections within texts	<p>Which sentence explains why (event) happened?</p> <p>a. The correct response is a reasonable statement of causation.</p> <p>b. Incorrect responses are events in the story that thoughtful reading reveals are not really the cause.</p>
Make connections among texts	<p>How does (story character's) feelings about (subject) compare to the poet's feeling about (subject)?</p> <p>a. The correct response identifies an appropriate similarity.</p> <p>b. Incorrect responses identify elements that exist in one passage but not in other.</p>

Source: From *Washington Assessment of Student Learning 4th-Grade Reading Test and Item Specifications*, 1998, Olympia, WA: Office of the Superintendent of Public Instruction. Reprinted by permission.

Decide if this amount of testing is acceptable. If the testing time as planned is unreasonable, modify your test by one of the following means:

- Convert some of the more time-consuming item formats to true/false items.
- Test only some of the learning targets—randomly select those to be covered, but be sure to tell students that this is what you will do.
- Combine learning targets and then report results or give descriptive feedback to students by target cluster rather than by learning target.

Assembling the Test

Begin the test with relatively easier items to maximize students' opportunity to start on a note of confidence. Consider arranging the items on your test according to the learning target each represents. If that presents a challenge in direction writing because you would be mixing item formats, consider indicating in some other fashion which learning target each addresses. However, do so only if this information will not give students unfair clues regarding the correct answer. Make sure your students know how many points each question is worth, so they can learn to prioritize their testing time. Put all parts of an item on the same page.

Stage 3: Critique the Assessment

There are two aspects of test quality to evaluate. The first is how well it matches the test plan and the second is how well the items are written.

Matching the Test Plan

To determine how well your actual test matches the test plan, you can use the process and form from Activity 4.4, "Analyze Your Own Assessment for Clear Targets."

Ensuring Item Quality

To evaluate the quality of the items you must do two things: make sure the item itself tests what you intended and check each item for how well it is written.

Item Tests What You Intended

If your items are sound in this respect, you can work backwards to turn them into the propositions you began with. Here's how:

- *Combine the multiple-choice item stem with the correct response.*
- *True true/false items already are propositions.*
- *Make false true/false items true to derive the proposition.*
- *Match up elements in matching exercises.*
- *Fill in the blanks of short answer items. (Stiggins, 2005, p. 109)*

If, by following these steps, you generate a list of the important learning targets or concepts you intended to test, then your items do, indeed test what you intended to test.

Item Is Well Written

You can audit each item using the information summarized in Figure 5.7 to determine if it is well written and to make adjustments as needed.

Stages 4 and 5: Administer the Assessment, Watch for Problems, and Revise as Needed

Even the best planning can't catch all problems with an assessment. Here are two things to watch for as you administer the test:

- Students have enough time to complete all test items. If students don't have the opportunity to attempt each item, their scores will not reflect what they have learned. Watch for students frantically marking items toward the end of the time allowed. Also look for large numbers of incomplete tests.
- Make notes on the questions for which students ask clarifying questions. Consider clarifying the directions or the item itself for the next time you use the test.

Figure 5.7 Test Item Quality Checklist

General guidelines for all formats

- Keep wording simple and focused. Aim for lowest possible reading level.
- Ask a question.
- Avoid providing clues within and between items.
- Correct answer should not be obvious without mastering material tested.
- Highlight critical words (e.g., most, least, except, not).

Guidelines for multiple-choice items

- State whole question in item stem.
- Eliminate repetition of material in response options.
- Be sure there is only one correct or best answer.
- Keep response options brief and parallel.
- Make all response options the same length.
- Limit use of all or none of the above.
- Use "always" and "never" with caution.

Guideline for true/false items

- Make them entirely true or entirely false as stated.

Guidelines for matching items

- Provide clear directions for the match to be made.
- Keep list of trigger items brief (maximum length is 10).
- Include only homogeneous items.
- Keep wording of response options brief and parallel.
- Provide more responses than trigger items.

Guideline for fill-in items

- Ask a question.
- Provide one blank per item.
- Do not make length of blank a clue.
- Put blank toward the end.

Figure 5.7 (Continued)

<p>Assembling the test</p> <p>_____ Arrange items from easy to hard.</p> <p>_____ Try to group items covering the same targets together, or identify the target each question addresses.</p> <p>_____ Try to group similar formats together.</p> <p>_____ Make sure the test is not too long for the time allowed.</p> <p>Writing directions</p> <p>_____ Write clear, explicit directions for each item type.</p> <p>_____ State the point value of each item type.</p> <p>_____ Indicate how the answer should be expressed (e.g., should the word true or false be written, or T or F? Should numbers be rounded to the nearest tenth? Should units such as months, meters, or grams be included in the answer?)</p> <p>Formatting test items</p> <p>_____ Be consistent in the presentation of an item type.</p> <p>_____ Keep all parts of a test question on one page.</p> <p>_____ Avoid crowding too many questions on one page.</p>

Source: Adapted from *Student-Involved Assessment for Learning*, 4th ed. (p. 107), by R. J. Stiggins, 2005, Upper Saddle River, NJ: Merrill/Prentice Hall. Copyright © 2005 by Pearson Education, Inc. Adapted by permission of Pearson Education, Inc.

TRY THIS**Activity 5.6 Critique the Items in Your Quiz**

Now return one final time to the 10-item quiz you developed in Activity 5.2 and evaluate the quality of each test item by working backward to generate the underlying propositions and checking each item against the checklist in Figure 5.7.

Selected Response Assessment *for Learning*

As we saw in Chapter 2, motivation to learn and level of achievement both rise when students are engaged in the assessment process. To see why this might occur, we have only to reflect on what happens to our own understanding of the intended learning when we as teachers engage in activities such as creating a test plan, writing propositions, and developing test items. The following ideas represent ways selected response tests can serve to help students answer the three essential questions at the heart of assessment *for learning*: “Where am I going?”; “Where am I now?”; and “How can I close the gap?” Assessment *for learning* strategies as applied to selected response tests are summarized in Figure 5.8.

Where Am I Going?

Strategy 1: Make Targets Clear

Once you have a test plan, you can use it, not just for test development, but also in service of assessment *for learning*. Give students a list of the learning targets, in terms they understand, at the beginning of the instructional time the test will cover. Let them develop a practice test plan based on their understanding of the relative importance of each learning target and then share your test plan. You can also use a test plan as a way to summarize the day’s or week’s learning by asking students to identify which cells your instruction has focused on.

Propositions can also play a role in assessment for learning. Explain to students that a proposition is a statement of important learning. Ask them to note what they understand to be the proposition(s) at the center of the day’s instruction. Have them keep a list and add to it each day. Help students see how well their lists match your own. Give groups of students your test plan and sample propositions representing the learning thus far. Have them match the propositions to the correct cell in the test plan.

Strategy 2: Use Strong and Weak Models

Give students an item formula, such as the ones in Table 5.4. (For younger students, you will need to translate the item formula into student-friendly language.) Show them a test item created with the formula. Ask them to identify which answers are wrong and which one is right by identifying the response pattern each follows. See Figure 5.9 for an example of how this can work when teaching fourth-grade students to infer.

Figure 5.8 Assessment for Learning Strategies: Assessing Knowledge and Reasoning Targets with Selected Response Tests

Where am I going?

1. Make targets clear.
 - Write targets in student-friendly language.
 - Share test plans at the outset.
 - Have students match propositions with test plan cells.
 - Have students develop propositions along the way.
2. Use strong and weak models.
 - Students identify wrong multiple-choice and fill-in answers and say why.

Where am I now?

3. Provide descriptive feedback.
 - Provide feedback target by target on a test.
 - Use definition of quality as basis for strengths and focus of improvement.
4. Teach students to self-assess and set goals.
 - Students use test plans as a basis for evaluation of strengths and areas for study.
 - Students complete self-evaluation and goal-setting form on basis of test or quiz results.

How can I close the gap?

5. Teach focused lessons.
 - Students use item formulas to write items.
 - Students answer question: *How do you know your answer is correct?*
 - Students turn propositions into items and practice answering the items.
 - Students create test items for each cell and quiz each other.
 - Students use graphic organizers to practice patterns of reasoning.
6. Students practice revising.
 - Students answer the question: *How do I make this answer better?*
7. Students reflect on and share what they know.
 - Students engage in self-reflection: I have become better at _____.
 - I used to _____, but now I _____.

Figure 5.9 Using Item Formulas with Students

Which one of these answers is a good inference, based on the reading selection from *The BFG*? Mark the good inference with a star. The right answer is a good inference because it is a guess based on clues from the story.

- a. The BFG could hear extremely well because he could not see very well.
- b. The author loved his father very much.
- c. The father did not finish high school.
- d. The father had a good imagination.
- e. The father wanted people to think he was a serious man.
- f. The father had a well-developed sense of hearing.
- g. The father was a funny person.

Some of these answers are wrong because they are not inferences at all! They are just facts that the story tells you outright. Write the letters of those wrong answers here:

Some of the answers are wrong because, even though they are guesses, there are no clues in the story to support them. Write the letters of those wrong answers here:

Be careful!!! You might think there is evidence for them, so look closely!!!

Where Am I Now?

Strategy 3: Provide Descriptive Feedback

To offer descriptive feedback, you must know what learning target each item addresses. Then you are able to couch your feedback in terms specific to that target. In the case of an inference test item, if a student chooses the answer that is wrong because, while it is an inference, it is not supported by evidence in the story, you can say, "You have selected an inference, so your answer does represent a guess, but the guess is not supported by evidence (or enough evidence) to be considered a good inference."

Strategy 4: Teach Students to Self-Assess and Set Goals

Hand your test plan out at the beginning of instruction. Have students self-assess on the learning targets or concepts as you teach them, using “traffic light” icons. Students mark the learning target or concept with a large dot—green to indicate confidence in having mastered it (“I’ve got it”), yellow to indicate a judgment of partial mastery (“I understand part of it, but not all of it”), or red to indicate little or no understanding (“I don’t get it at all”). Then let the “greens” and “yellows” partner to fine tune their understanding while you work with the “reds.” (Black, Harrison, Lee, Marshall, & William, 2002).

A few days prior to the test, have students self-assess on the learning targets or concepts represented on your test plan. Then have them create a study plan based on what they think they have yet to learn. Or, have students assign themselves to study groups based on their assessment of what they need to learn.

Because students often take tests without knowing what the test measures beyond the most general level (“reading,” “social studies,” “science”), when they are asked to use test results to share what they know or to set goals for future learning, what they produce is often too general to be of much use: “I am pretty good at math”; “I can read better than I used to”; “I need to study more.” Although “take my book home” or “try harder” are noble thoughts, they do not reflect either what students have learned or what they may actually need to learn, and therefore are of limited use. Activity 5.7, beginning on page 158, is designed to help students accurately identify what they know and set goals for their next steps.

How Can I Close the Gap?

Strategy 5: Teach Focused Lessons

Let students create test items using an item formula. See Figure 5.10 for an inference example that would work well as a followup to the activity in Figure 5.9.

After each multiple-choice item on a test, ask students to explain the reason for their selection by including the question, “How do you know your answer is correct?” and providing several lines for a response. Discuss common reasons for right and wrong choices when you pass back the test. Figure 5.11 provides an example with a question asking about an author’s purpose.

Assign groups of students to each cell of your test plan. Have them create questions that might be on the test, based on the propositions they generated for each cell during instruction. Have the groups take each others’ practice tests.

Teach students to use graphic organizers as a means to understand the specific kind of reasoning called for. After doing activities from Strategies 1 and 2, let students create their own graphic organizer for a specific pattern of reasoning.

Figure 5.10 Student-Generated Inference Questions

How to Write an Inference Question

Here is a recipe that test writers use to create multiple-choice inference questions and answers.

Question:

Which idea does this selection suggest?

Possible responses (*these include the right answer and several wrong answers*):

- a. The correct response is a guess that is supported by clues in the text.
- b. One incorrect response is a wild guess, because there aren't any clues in the text to support it.
- c. Another incorrect response is just a detail from the text recopied. It's not a guess at all.

Now it's your turn!

First, read the assigned passage in the text. Then, work with a partner to create the right and wrong answers to the inference question below.

Here's the inference question:

Which idea does this selection suggest?

You write a correct answer and two incorrect answers. You can mix up the order—the correct answer doesn't have to come first.

- a.
- b.
- c.

Figure 5.11 Example of Followup Question to Probe Reasoning

Which of the following BEST sums up the author's purpose in writing this passage?

- a. To show the character of Danny's father.
- b. To tell a funny story about a giant.
- c. To explain where dreams come from.
- d. To keep children from being afraid.

How do you know your answer is correct? _____

Strategy 6: Students Practice Revising

Anything we do to give students practice with applying what they know about quality or correctness to rework their own answers or to offer suggestions on someone else's work causes them to engage in revision. Consider letting them first practice on anonymous responses by answering one or both of the following questions: "What is wrong with this answer?" "What would make this answer better?"

Strategy 7: Students Reflect on and Share What They Know

We may think of this strategy as best suited to performance assessment, but it is equally effective with selected response assessment. Students should be thinking about their achievement with respect to the knowledge and reasoning targets measured by this method, both because they play an important role in their education, and also because it is these fundamental targets that struggling students often have not mastered. As we stated in Chapter 2, any activity that requires students to reflect on what they are learning and to share their progress both reinforces the learning and helps them develop insight into themselves as learners. These are keys to enhancing student motivation.

Some software programs used for reading instruction and assessment harness the power of Strategy 7. They have built-in mechanisms for students to track their progress and communicate their results on reading comprehension learning targets. Students read text varying in

length from a passage to a book, answer a variety of questions, and get immediate feedback on how they did. They are able to monitor their own progress and experience the joy of watching themselves grow. For many students, that joy causes them to like the assessment experience, even if they are not wildly successful at first, and it also motivates them to continue reading.

TRY THIS

Activity 5.7 Engaging Students in Self-Reflection and Goal Setting with Selected Response Tests

Choose a selected response test you intend to give to students and complete the following steps. (Blank forms for both an elementary and secondary version of this activity can be found on the CD in the file, "Goal Setting with Tests.")

1. Make a numbered list of the learning targets represented on the test. Figure 5.12 shows an example from a ninth-grade biology class.
2. Transfer that information to the chart in Figure 5.13a by filling out the "Learning Target #" column, which identifies the learning target addressed by each item. Copy the chart for each student and hand it out with the test. Figure 5.13b shows adaptations made for the ninth-grade biology example in Figure 5.12.
3. As students take the test, they note on the chart whether they feel confident or unsure of the correct response to each item. Correct the tests as usual and hand them back, along with the numbered list of learning targets. The students are now ready to identify their own specific strengths and areas for further study by following the steps explained in Figure 5.14.

This activity is most powerful as a learning experience if students have an opportunity to take some version of the test again, as when studying for a retake of the test, or after a quiz or a practice test in preparation for a final exam. In the case of a retake, you may want students to follow up their self-analysis with a specific study plan. Figure 5.15 shows examples of forms used for that purpose.

Figure 5.12 List of Learning Targets Tested—Ninth-Grade Biology Example

Below are 14 Key Learning Targets for Unit 9a. In the student assessment of your achievement of these learning targets, you will identify the areas in which you demonstrated proficiency and the areas in which you need to do additional study and preparation for mastery of the Unit 9 skills and knowledge.

1. Recognize that ecology is the scientific study of the interactions between organisms and their environment
2. Distinguish between a population, community, ecosystem, biome and biosphere
3. Describe how organisms interact with each other in different ways (producers, consumers, predator, prey, scavengers, parasites, decomposers) to transfer energy and matter in an ecosystem
4. Recognize that energy flows from one trophic level (one direction only) to another
5. Recognize 90% of the energy of a trophic level is lost during life processes and as heat in the transfer to the next trophic level
6. Describe how energy relationships can be represented and calculated in food/energy, biomass and numbers pyramids
7. Explain all energy for an ecosystem originates from the sun
8. Diagram the relative amounts of energy in a trophic level using an ecological pyramid
9. Diagram the flow of energy in a food chain or food web
10. Explain why matter is constantly recycled in an ecosystem
11. Recognize each element is cycled in a specific way
12. Express how the recycling of matter is necessary to make it available for organisms to use
13. Distinguish between the four biogeochemical cycles (H_2O , CO_2 , O_2 , and N_2)
14. Explain the major steps in each of the four biogeochemical cycles (H_2O , CO_2 , O_2 , and N_2)

Source: Steve Wavra, Sweetwater Union High School District, Chula Vista, CA. Used by permission.

Figure 5.13a Student Documentation of Selected Response Test Results

Identifying Your Strengths and Focusing Further Study

As you answer each question on the test, decide whether you feel confident in your answer or are unsure about it, and mark the corresponding box.

Problem #	Learning Target #	Confident	Unsure		Right	Wrong	Simple Mistake	Further Study

- 1. After your test has been corrected, identify which problems you got right and which you got wrong by putting Xs in the "Right" and "Wrong" columns.
- 2. Of the problems you got wrong, decide which ones were due to simple mistakes and mark the "Simple Mistake" column.
- 3. For all of the remaining wrong answers, mark the "Further Study" column.

Source: Adapted from *Assessment FOR Learning: An Action Guide for School Leaders* (p. 198), by S. Chappuis, R. J. Stiggins, J. Arter, & J. Chappuis, 2004, Portland, OR: Assessment Training Institute. Copyright © 2006, 2004 Educational Testing Service. Adapted by permission.

**Figure 5.13b Student Documentation of Selected Response Test Results—
Biology Teacher’s Variation**

Identifying Your Strengths and Focusing Further Study

As you answer each question on the test, decide whether you feel confident in your answer or are unsure about it, and mark the corresponding box.

Question	Key Learning	Knew It	Wasn't Sure	Guessed	Got It Right	Got It Wrong	Simple Mistake	Misread Question	Need to Re-Study
1	1								
2	2								
3	3								
4	7								
5	3								
6	3								
7	3								
8	3								
9	9								
10	6								
11	2, 9								
12	3								
13	4, 6, 9								
14	3, 4								
15	3, 4								
16	6								
17	6								
18	5								
19	5								
20	6								
21	10, 11, 12								
22	13								
23	5, 11								
24	11, 12, 13								
25	11, 12, 13								
26	11, 13								

Source: Steve Wavra, Sweetwater Union High School District, Chula Vista, CA. Used by permission.

Figure 5.14 Student Analysis of Selected Response Test Results

1. To identify your areas of strength, write down the learning target numbers corresponding to the problems you felt confident about *and* got right. Then write a short description of the target or problem.

MY STRENGTHS:

Learning Target #	Learning Target or Problem Description

2. To determine what you need to study most, write down the learning target numbers corresponding to the marks in the "Further Study" column (problems you got wrong, NOT because of a simple mistake). Then write a short description of the target or problem.

MY HIGHEST PRIORITY FOR STUDYING:

Learning Target #	Learning Target or Problem Description

3. Do the same thing for the problems you were unsure of and for the problems on which you made simple mistakes.

WHAT I NEED TO REVIEW:

Learning Target #	Learning Target or Problem Description

Source: Adapted from *Assessment FOR learning: An Action Guide for School Leaders* (p. 199), by S. Chappuis, R. J. Stiggins, J. Arter, & J. Chappuis, 2004, Portland, OR: Assessment Training Institute. Copyright © 2006, 2004 Educational Testing Service. Adapted by permission.

Figure 5.15 Student Goal-Setting Frames

To get better at _____, I could . . .

-
-
-
-
-

Some things I am going to start doing are . . .

-
-
-

I'll start doing this on _____ and work on it until _____.

(date) (date)

One way I'll know I'm getting better is . . .

Goal	Steps	Evidence
What do I need to get better at?	How do I plan to do this?	What evidence will show I've achieved my goal?
<p><i>Time Frame:</i> Begin _____ End _____</p> <p>Date _____ Signed _____</p>		

Source: From *Self-Assessment and Goal-Setting*, (p. 45) by K. Gregory, C. Cameron, & A. Davies, 2000, Merville, BC: Connections. Reprinted by permission.

Summary

In this chapter we revisited the idea that selected response items are a good way to measure knowledge and reasoning learning targets, as long as students can read at the needed level to understand the questions. Selected response is an efficient way to cover a lot of content in a short period of time.

Although we reviewed all steps in test development—from planning to critiquing the final product—we focused on creating test plans, generating propositions to identify important content, and adhering to guidelines for writing high-quality selected response items of all types.

Finally, we offered concrete examples of how to involve students in selected response assessment. These strategies focused on how to use these tests to help students answer the three questions that define assessment *for* learning: “Where am I going?”; “Where am I now?”; and “How can I close the gap?”

■ Tracking Your Learning—Possible Portfolio Entries

Any of the activities from this chapter can be used as portfolio entries. Remember, the learning targets for this book are outlined in Figure 5.1, listed in Table 1.2, and described in detail in Chapter 1. The portfolio entry cover sheet provided on the CD in the file, “Portfolio Entry Cover Sheet,” will prompt you to think about how each item you select reflects your learning with respect to one or more of these learning targets.

Any of the following activities would make good portfolio entries. Each could also be used as a learning team activity.

TRY THIS

Activity 5.8 Create a Selected Response Assessment

Select an upcoming context (a unit, an area of study, a collection of related learning targets) for which a selected response test would be the appropriate assessment and create that assessment, following the guidelines described in this chapter.