3

Formats for Teacher-Made Tests

Teacher-made tests and other assessments should have as their primary goal the improvement of learning. Good assessments show teachers whether students truly understand the learning and provide insight into the student's ability to use levels of thinking. Assessment should be on-going, not something tacked on at the end of a unit of study. Assessment should be frequent enough to let the teacher and students know if they are making good progress, and they should be consistent in content application and grading. Teachers make assessments daily through informal observations, the gathering of data, and the more formal classroom test.

There are seven basic types of assessment questions or formats for teacher-made tests. Some authors divide the formats differently, but all authors use the same basic formats. In the following section, I'll provide a short definition of each type and indicate the strengths and weaknesses of each.

FORCED-CHOICE QUESTIONS

Rick Stiggins (1994) defines forced-choice assessment as

the classic objectively scored paper-and-pencil test. The respondent is asked a series of questions, each of which is accompanied by a range of alternative responses. The respondent's task is to select either the correct or the best answer from among the options. The index of achievement is the number or proportion of questions answered correctly.

Examples of the forced choice are conventional multiple-choice, matching, alternate-choice, true/false, multiple-response, and fill-in-the-blank questions. Forced-choice questions were found to be effective to a low degree when assessing for processes, communication skills, and on nonachievement factors (such as the ability to get along with others). They had only a moderate impact on assessing declarative knowledge and thinking and reasoning skills.

Forced choice is the most commonly used format for standardized tests and is one of the poorest ways to determine whether students can use the information learned. This group of tests is called "forced" because there is generally only one right answer, so respondents are constrained by the choices offered (Stiggens, 1994). These assessments should be given when teachers want to test declarative knowledge, such as facts, dates, formulas, and general information on a topic. The next sections look at six of the most common forms of questions for forced-choice assessments.

Conventional Multiple Choice

The conventional multiple-choice format contains a *stem*, a number of *distractors*, and one correct choice. For example,

The best definition for evaluation is . . . [*stem*]

A. The collection of student data. [*distractor*]

B. The process of making a judgment. [*correct choice*]

C. The process of monitoring. [*distractor*]

According to Marzano (2000),

As tools for classroom assessment, multiple-choice items are fairly difficult and time-consuming to write. Probably the most difficult aspect of writing multiple-choice items is designing viable distracters. They must be inaccurate enough to be considered wrong by students who understand the content, but plausible enough to be selected by students who are making an educated guess.

Matching

In tests using matching questions, students are instructed to match a given item with the option that best fits. For example,

For each item below, select the option that accurately completes the statement.

Item	Options
A. $5 \times 12 =$ _____	1. 144
B. $3 \times 11 =$ _____	2. 60
C. $8 \times 12 =$ _____	3. 222
D. $6 \times 11 =$ _____	4. 66
	5. 96
	6. 31
	7. 26
	8. 33

The main advantage to this type of format for teachers is that it takes time to create, since the correct choice for one item is a distracter for the other items. Moreover, because there are more choices than answers, students cannot use the process of elimination to find the correct answer.

Alternate Choice

Alternate-choice test questions are like multiple-choice questions, except there are only two answer options. For example,

____ 1. An (a. *assessment*, b. *evaluation*) is a process for making a judgment.

____ 2. (a. *Formative assessment*, b. *Summative assessment*) occurs before and during instruction.

This format works when the teacher wants to test for knowledge of terms and phrases. One of the biggest disadvantages is that students have a 50% chance of guessing the correct answer.

True/False

In the true/false format, students must identify a statement as accurate (*true*) or inaccurate (*false*). For example,

____ 1. Franklin D. Roosevelt was a war hero.

____ 2. Winston Churchill was a war correspondent.

____ 3. Franklin D. Roosevelt coined the phrase "V for victory."

Although this type of test gives the appearance of being easy to write, good true/false questions are actually difficult to construct. Writing false statements is much like composing distractors in multiple-choice items: They must be plausible enough to attract those students who do not really know the answer. And, like the alternate-choice questions, students have a 50% chance of guessing the correct answer.

As a word of caution to the new teacher, let me note that students may make their T's and F's look alike so that no matter what the answer choice, they can say that they wrote

the correct answer. You can prevent this by having students write out the words *true* or *false*, or by having them use O's for true and X's for false.

Multiple Response

Multiple-response questions allows for more than one correct answer. For example,

Meg Byers is a first year teacher at South Middle School, where she teaches reading. She has noticed that although her students can read the vocabulary in print, they often read without understanding. Which of the following teaching tools will be helpful to her as she works with her students to improve comprehension?

1. Help students connect the reading to something they already know.

2. Have students read the information in small chunks followed by questions about what they read.

3. Keep them after school if they do not understand their work.

4. Demonstrate to students how to use self-talk as they read.
   A. 1, 2, 3, and 4
   B. 1, 2, and 4 only
   C. 1, 3, and 4 only
   D. 2 and 3 only

A variation of the multiple-response question might be,

Which of the following types of test questions has a stronger impact on assessing procedural knowledge? Place a *Y* in front of the answer(s) if they have a high effect and an *N* in front of the answer(s) that do not have a high effect on procedural knowledge.

____ Forced answer

____ Essay

____ Self-assessment

____ Performance task

____ Short written response

____ Oral reports

The advantage of this type of test is that it allows the teacher to test more than one feature of the student's knowledge. For example, in the example above, students must know the impact on assessment of all of the examples given.

Fill in the Blank

In fill-in-the-blank items, students must supply the missing word or words that make sense within a sentence. For example,

The most important function of testing is _____.

This is the easiest type of test to create, because the stem is a sentence or phrase that must be completed.

ESSAYS

Essay questions are believed to be one of the oldest forms of questions used on teacher-made tests. The Website of the National Center for Research on Evaluation, Standards, and Student Testing (CRESST, www.cresst.org) provides the following example for a written essay:

Since the start of the year, your class has been studying the principles and procedures used in chemical analysis. One of your friends has missed several weeks of class

because of illness and is worried about a major exam in chemistry that will be given in two weeks. This friend asks you to explain everything that she will need to know for the exam.

Write an essay in which you explain the most important ideas and principles that your friend should understand. In your essay, you should include general concepts and specific facts you know about chemistry, and especially what you know about chemical analysis or identifying unknown substances. You should also explain how the teacher's demonstration illustrates important principles of chemistry. Be sure to show the relationships among the ideas, facts, and procedures you know.

The strength of essay-type assessments is in the stem for the essay itself. What do you want to know? Do you want to know if students understand the facts or do you want to know if they can use reasoning, problem-solving techniques, or decision making? The stem should clearly indicate what it is that you want to assess. A rubric should be provided for essay questions so that the students are clear about what it is they are being tested over.

Short Written Responses

Short written responses are really mini-essays in which students give a short answer to a given question. These types of questions are used when the teacher wants to know if the students understand the information. Some use of higher-level thinking skills may also be seen in this type of question, although it is limited somewhat by the length of the answer.

Skillfully written questions can probe beyond basic knowledge and comprehension. For example, in the short story "After Twenty Years," by O'Henry, two friends make a pact to meet 20 years after graduation to see what has happened in their lives. Jimmy stays in the area and becomes a

police officer while his best friend takes off to other parts of the country and gets in trouble with the law. On the fateful night in which they are to meet, Jimmy starts down the street only to see his friend under a lamppost and recognizes him from a wanted poster. Jimmy must decide if he will arrest his friend or not. A possible short written response that would require students to think at a high level might be, "Did Jimmy use good judgment in his decisions about his friend? Why or why not? Defend your answer."

Short-Answer Questions

Short-answer questions are also mini-essays. The basic difference between a short answer question and a short response is length: A short response might be a few words or even one word to complete a prompt, whereas a short-answer question is written in whole sentence form and tends to be longer in length than short response. An example of a short-answer question at the elementary level might be, "In *The Math Curse*, by Sczescka, what are some ways the young girl was cursed by math?" At the secondary level, an example would be, "What was the major impact of the Boston Tea Party?"

These questions are best used when a teacher wants to know if students understand declarative knowledge or informational topics. These test questions are limited by the shortness of the response. Short-answer questions have a low impact on assessing processes, communication skills, and nonachievement factors. They have a moderate impact on assessing thinking and reasoning skills, and a high impact on assessing informational topics (declarative knowledge).

ORAL REPORTS

Oral reports are essays that are presented out loud. They have

they assess students' speaking ability. They are highly effective at assessing informational topics and thinking, reasoning, and communication skills. Oral reports have a moderate impact on testing process topics and a low impact on assessing nonachievement factors.

PERFORMANCE TASKS

Performance tasks require students to show what they can do with the information learned. *Performance task* is sometimes used synonymously with *authentic task*, but assessment purists will argue that performance tasks are contrived to determine whether students can use the information learned, whereas authentic tasks mirror the real world. For our purposes, the term *performance task* includes both ideas. Even a contrived task to determine if students can use information can be discussed in terms of how it might be done in the real world.

Performance tasks are ranked high in assessing accurately four of the five categories: assessing informational topics, process topics, thinking and reasoning skills, and communication skills. Performance tasks can be used to assess nonachievement factors but have a moderate impact.

TEACHER OBSERVATION

Teacher observation is an informal type of assessment used to examine a type of behavior, such as the ability to work with others. It can also be used to observe process skills, such as a student's ability to use a graph. Teacher interviews of students also come under this classification.

Teacher observation has high impact on the assessment of process topics and nonachievement factors, a moderate impact on testing informational topics, and a low impact in assessing thinking and reasoning as well as communication skills

STUDENT SELF-ASSESSMENT

Student self-assessment is the most underused form of classroom assessment. Students assess their own thinking, work, processes, and learning in this format. Student self-assessment has a high reliability in all five of the assessment categories: informational topics, process topics, thinking and reasoning skills, communication skills, and nonachievement.