CHAPTER 17

Writing Multiple-Choice and Other Objective Tests

Some valuable ideas you'll find in this chapter

- **Test blueprints – lists of the learning goals covered by the test – are vital to planning effective multiple-choice tests.**
- **Well-designed multiple-choice questions give useful diagnostic information on where the student's thinking went wrong.**
- **Multiple-choice and matching questions can assess some kinds of thinking skills, including the abilities to apply and to analyze.**

While performance assessments (Chapter 16) are growing in popularity, multiple-choice and other objective tests may still have a place in some assessment toolboxes. This chapter discusses how to plan an objective test, how to write effective multiple-choice and other objective test items, and how objective tests can be used to assess some thinking skills.

What is an objective assessment?

Jargon Alert!

Objective and Subjective Assessments

An *objective* assessment has only one correct answer for each question, while in a *subjective* assessment a variety of answers may be correct.

Jargon Alert!

Quantitative Assessments

Quantitative assessments use structured, predetermined response options that can be summarized into meaningful numbers and analyzed statistically. Multiple-choice tests, rubrics (Chapter 15), and rating scales (Chapter 20) are examples of quantitative assessments.

Subjective assessments (see Jargon Alert) require professional judgement to score. *Objective* assessments (see Jargon Alert) can be scored accurately by a reasonably competent eight-year-old armed with an answer key, although *interpretation* of the scores requires professional judgment. Objective assessments are almost always tests, while subjective assessments may be tests or performance assessments (Chapter 16).

Some people confuse *quantitative* assessments (see Jargon Alert) and *objective* assessments, erroneously thinking that quantitative assessments are objective. Actually, it's the other way around. Objective tests are quantitative, but so are subjective assessments using rubrics and rating scales.

Why use an objective test?

Objective tests remain widely used for several reasons.

Students can provide a great deal of information on a broad range of learning goals in a relatively short time. Testing experts call this *efficiency*. If you want to assess student learning on a wide array of concepts and skills, a 45-minute multiple-choice test will give you more comprehensive information on their learning than a 45-minute essay test.

Objective tests encourage broader – albeit shallower – learning than subjective assessments because of their efficiency. Asking students to write a paper on a poem by Wordsworth is a good choice if the learning goal is to analyze that poem

thoroughly but a poor choice if the learning goal is to analyze Romantic literature. For the latter goal, an objective test asking students to react to a variety of examples of Romantic literature might be a better choice.

Well-constructed multiple-choice tests can help diagnose problem areas. Consider this simple example:

What is 2 × .10?

A. 20
B. 2.10
C. .2
D. .02

Each distracter gives a clue on where the student's thinking goes wrong. Choosing B, for example, indicates that the student confuses multiplication and addition signs. Choosing A indicates that the student confuses multiplication with division, while choosing D indicates a problem with decimal place value. Faculty and staff can thus use the evidence of student learning from well-written multiple-choice tests to identify areas of difficulty and help their students accordingly.

Objective tests are fast and easy to score, although they are difficult and time-consuming to construct. If they are stored securely so they can be reused, the payback on the time spent writing them increases.

Some people think that objective tests are more valid or accurate than subjective assessments. To the contrary, both objective and subjective assessments require professional judgment. In objective tests, professional judgment goes into developing or choosing test questions, which can vary considerably in their difficulty and in the learning goals they assess, and in setting standards for adequate performance (Chapter 22). So *all* assessments are subjective, in that they all require professional judgment, and objective assessments are not necessarily better than subjective ones.

Why not use an objective test?

Just like any other assessment, objective tests have shortcomings and are not always appropriate. Subjective assessments such as performance assessments are increasingly popular for several reasons.

Objective tests cannot assess many important thinking skills, including organization, synthesis, and original thinking.

Objective tests are less convincing evidence of real-world skills than authentic assessments (Chapter 16). Writing samples, for example, are more convincing evidence of writing skill than answers to multiple-choice questions on how to write. Similarly, watching dental students treat a cavity provides more compelling evidence of their skill than answers to multiple-choice questions on how to do this procedure. (Assessing performances such as this procedure is discussed in Chapter 22.)

Objective test scoring procedures usually don't allow for nuances. On a subjective math test, for example, students can receive partial credit for doing part of a multi-step problem correctly. But on a multiple-choice math test, they receive no credit for an incorrect answer, even if they do much of their work correctly.

Objective test items are hard to write. Writing good, clear multiple-choice items with good distracters can be difficult and time-consuming.

Objective tests sometimes require significant reading ability. If so, they may not fairly assess students who have achieved the test's learning goals but read slowly or for whom English is a second language. While reading is an essential skill, if we are assessing understanding of science concepts, we want an assessment tool that assesses just that.

Jargon Alert!

Test-Wise

Test-wise students are adept at discerning inadvertent clues to correct and incorrect multiple-choice options. This increases their chances of choosing the correct answer, even if they haven't learned what the test is assessing.

It's possible to guess the correct answer to most objective items through either plain luck or test-wise skills (see Jargon Alert). While a well-constructed test minimizes test-wise students' advantage, there is always the possibility that students who haven't learned what they're supposed to will do relatively well on a test through chance alone.

Jargon Alert!

Test Blueprint

A *test blueprint* or *table of specifications* is list of the learning goals addressed on the test. It may also list the number of test items assessing each learning goal.

Planning an objective test

Chapter 16 explained that effective assignments are planned by developing a rubric: A list of the learning goals or traits that students are to demonstrate in the completed assignment. Effective tests are similarly planned by developing a test blueprint (see Jargon Alert). Exhibit 17.1 is an example of a test blueprint for an exam in an Educational Research Methods course.

Exhibit 17.1 A Test Blueprint for an Exam in an Educational Research Methods Course

(There is one test question on each topic unless otherwise indicated.)

Sampling

- ☐ Recognize the difference between a sample and a population.
- ☐ Understand how each type of sample is selected.
- ☐ Choose an appropriate sample size.

Instrumentation and Survey Research

- ☐ Identify the relative merits and limitations of published and locally developed instruments.
- ☐ Recognize examples of each of the four frames of reference for interpreting scores.
- ☐ Recognize appropriate uses of each item format (such as multiple-choice and Likert scale).
- ☐ Understand the characteristics of a good instrument item, including how to avoid biased questions.

Descriptive Statistics

- ☐ Select the most appropriate descriptive statistic for a given research situation.
- ☐ Use percentage guidelines to interpret standard deviations.
- ☐ Identify the direction and strength of r and/or a scatterplot.
- ☐ Identify the likely direction and strength of a correlation between two given variables.

Validity and Reliability

- ☐ Identify the type of reliability or validity evidence provided by given information on an instrument. (2 questions)
- ☐ Understand the meaning and implications of measurement error.
- ☐ Recognize examples of measurement error.
- ☐ Understand the general principles for ensuring validity.

Inferential Statistics

- ☐ Select the most appropriate inferential statistic (t, F, or χ^2) for a given research situation. (2 questions)
- ☐ Identify the most common cutoff points that statisticians use in deciding whether two means differ statistically significantly from one another.
- ☐ Interpret the results of t-tests as presented in research articles.

Experimental Research

- ☐ Interpret the symbolic representations of experimental designs.
- ☐ Identify the appropriate research design for a given research situation.

Correlational Research

- ☐ Explain what r^2, R, R^2, and partial correlations are and what they tell us. (2 questions)
- ☐ Explain what regression analysis is used for and what it tells us.

Why create a test blueprint?

Chapters 5 and 16 note that *students spend their learning time and energies focusing on what they're graded on*. No matter what your stated learning goals are, if your test focuses on memorized knowledge, for example, that's what students will spend their time learning, and that's what they'll take away from the learning experiences you

provide them. Test blueprints help ensure that students focus on learning what you think is most important. They accomplish this in several ways:

Test blueprints help ensure that the test focuses on the learning goals you think are most vital. Suppose that you are writing a test for Units 8, 9, and 10 of a course. While you consider Unit 10 the most important of the three, you may find that it's much easier to think of test questions for Unit 8. If you write a test without a blueprint, you can easily end up with too many questions on Unit 8 and too few on Unit 10. Students taking such a test may be able to earn a decent score without having mastered important learning goals of Unit 10.

Test blueprints help ensure that a test gives appropriate emphasis to thinking skills. Faculty writing test questions without a blueprint often end up with questions asking for basic conceptual knowledge rather than questions asking students to interpret, infer, analyze, or think in other ways. In fact, tests written without blueprints sometimes become tests of trivia rather than tests of thinking skills. Students who do well on such tests may not have mastered important skills, while students who have truly learned those important thinking skills may nonetheless earn low scores.

Test blueprints make writing test questions easier. Armed with a test blueprint, you'll know exactly what must be covered on the test (one question on Concept A, two on Skill B, and so on), and you'll spend less time pondering what questions to write.

Test blueprints help document that students have achieved key learning goals. As discussed in Chapter 23, scores on each test item, accompanied by test blueprints that describe each item's learning goal, are direct evidence (Chapter 3) of exactly what students have learned.

Identify the learning goals you'd like to address on the test

Start creating a test blueprint by using syllabi, lesson plans, learning activities, and other curricular materials to list the learning goals that you'd like to assess. Phrase your learning goals using action verbs that describe what students should be able to think and do (Chapter 4). Instead of simply listing *Hemingway* in a test blueprint on twentieth-century American literature, for example, state the knowledge and skills you want students to demonstrate regarding Hemingway (perhaps *Identify works written by Hemingway* and *Distinguish Hemingway's writing style from those of his peers*). Keep in mind that objective tests can assess more than basic knowledge and understanding – you'll see how later in this chapter.

Decide how many test items you'd like for each learning goal

You may decide that some learning goals are so important that you want three, four, or more questions on each of them, to be sure that students have achieved the goal and can address it from multiple perspectives or contexts. Other learning goals may need only one question.

For a comprehensive exam, decide how many test items you'd like on each major unit

A midterm exam, for example, might cover the first five units in the course syllabus. Decide how many test questions you want on each of those units, in proportion to the relative importance of each unit (perhaps 6 questions on Unit 1, 10 questions on Unit 2, and so on).

Finalize the test blueprint and provide it as a study guide

Compare the number of items you'd like to have on each learning goal or objective (and, for a comprehensive exam, on each major unit) with the total number of items you can realistically expect students to answer in the testing period. You may find that you can't address every learning goal on the test. Perhaps, for example, you'd need 40 items on the test to cover your learning goals, but students will be able to answer only 30 in the testing period. You'll need to eliminate some of the less-important learning goals to make sure the test gives enough attention to the most important ones.

Once you have finalized the test blueprint, provide it to your students to help them focus their studies on the learning goals that you think are most important.

Decide how difficult your test should be

Objective tests have two different assessment purposes. One is to make sure that students have learned the essentials – matters so vital that the consequences of not having learned them are dire. In health and medical disciplines, for example, if students haven't mastered every key learning goal, a future patient might conceivably die. In other disciplines such as laboratory science and culinary arts, if students haven't mastered key learning goals on safety, someone might be gravely injured. And in many disciplines, certain basic concepts or skills must be mastered in order to succeed in later courses or in a career. In these situations, students who have studied and prepared well should be expected to answer virtually every item correctly. A score below 90 percent or even 95 percent for these kinds of test items might be considered failing.

The other potential purpose of an objective test is to distinguish students with superb levels of learning from those with mediocre levels of learning – to identify the students who deserve an A and not just a passing grade. These tests have challenging questions that only the top students answer correctly.

Many tests aim to serve both purposes. Some test items address fundamental essentials that virtually every student should answer correctly, while other test items are intentionally difficult to challenge the very best students and thereby separate the A, B, and C students.

Review your test blueprint and identify which learning goals are essential and which are more challenging. Writing difficult multiple-choice questions to separate the A, B, and C students is, well, remarkably difficult. It's very hard to come up with a test question that assesses an important learning goal but that a large proportion of students will get wrong. It's much easier to write difficult items that are on minutiae, trick questions that require nuanced reading, or questions that assess logical reasoning skill rather than key learning goals. Follow the suggestions in List 17.1 to write challenging multiple-choice questions that assess important learning goals.

List 17.1 Tips for Writing Challenging Rather Than Trick Questions

- ☐ **Use a test blueprint.** It ensures that each item assesses an important learning goal.
- ☐ **Make your tests open-book, open-note.** Tell students they can use any resource they like except a friend or the means to communicate with one. Using open-book, open-note tests forces you to eliminate items assessing simple knowledge that students can look up. Your test will include only items that assess deeper comprehension and thinking skills.
- ☐ **Build items around common misconceptions.** Many people, for example, think that plants get nutrients only from soil and water, not air; this misconception can become the basis of an effective botany test question.
- ☐ **Create interpretive exercises** (discussed later in this chapter). They assess thinking skills such as application and analysis.
- ☐ **Evaluate your test results** using the tools in Chapter 24. Revise any unnecessarily difficult items before including them in another test.

Writing good multiple-choice items

As with any other assessment, multiple-choice tests should yield fair and truthful information on what students have learned (Chapter 3). List 17.2 offers suggestions for writing fair and truthful multiple-choice items (Paniagua & Swygert, 2016; Parkes & Zimmaro, 2016; Haladyna, 2015) that follow two basic precepts.

- **Remove all the barriers that will keep a knowledgeable student from answering the item correctly.** Students who have *truly* learned the concept or skill that an item assesses should choose the *correct* answer.

- **Remove all clues that will help a less-than-knowledgeable student answer the item correctly.** Students who truly *haven't* learned the concept or skill that an item assesses should answer the item *incorrectly*.

List 17.2 Tips for Writing Good Multiple-Choice Questions

General Tips

- ☐ **Keep each item as concise as possible.** Short, straightforward items are usually easier for students to understand than complex statements. Avoid irrelevant material, digressions, and qualifying information unless you are specifically assessing the skill of identifying needed information. Don't repeat the same words over and over in the options; put them in the stem (see Jargon Alert).
- ☐ **Define all terms carefully.** What do you mean by *sometimes*, *usually*, or *regularly*? If you ask *Which bird is largest?* make clear whether you mean largest in terms of wingspan or weight.
- ☐ **Don't make the vocabulary unnecessarily difficult.** Except for terms you are specifically assessing, keep your vocabulary simple – perhaps high school level. Otherwise, you may unfairly penalize students who have achieved your learning goals but don't have a strong general vocabulary.
- ☐ **Watch out for interlocking items** in which a student can discern the answer to one from the content of another. Review carefully all items that share similar options. Also don't ask students to use their answer to one question to answer another. If they get the first question wrong, they will automatically get the other question wrong as well, even if they have learned the concept or skill assessed in the second question.

Writing a Good Stem

- ☐ **The stem should ask a complete question.** The student shouldn't have to read the options to discern the question. To check this, see if students would be able to answer the question posed in the stem correctly if no options were provided.
- ☐ **Avoid "Which of the following" items.** They require students to read every option and can penalize slow readers in a timed testing situation.
- ☐ **Ask questions that can't be answered from common knowledge.** Someone who hasn't studied the material shouldn't be able to answer the questions correctly.
- ☐ **Avoid negative items.** In a stressful testing situation, students can miss the word *not* or *no*. If you must have negative items, underline, capitalize, and/or boldface words like **NOT** or **EXCEPT**.
- ☐ **Avoid grammatical clues to the correct answer.** Test-wise students know that grammatically incorrect options are wrong. Use expressions like *a/an*, *is/are*, or *cause(s)*.

Writing Good Options

- ☐ **You needn't have the same number of options for every question.** Three options are fine (Rodriguez, 2005), and some questions may have only three plausible options (such as *Increases, Decreases, Remains unchanged*). Only rarely are more than four options needed. A good fifth option is often hard to come up with, takes extra reading time, and only reduces the chances of randomly guessing the correct answer from 25 percent to 20 percent.
- ☐ **Order options logically.** Order options numerically if they are numbers, and alphabetically if they are single words. This helps students who know the answer find it quickly. If the options have no intuitive order, insert the correct answer into the options randomly.

(continued)

List 17.2 Continued

- ☐ **Line up options vertically rather than horizontally.** It's much easier – and less confusing – to scan down a column than across a line to find the correct answer. If you are using a paper test and your options are so short that this seems to waste paper, arrange your test in two columns.
- ☐ **Make all options roughly the same length.** Test-wise students know that the longest option is often the properly qualified, correct one.

Jargon Alert!

Stems, Options, Responses, Distracters, and Foils: The Elements of Multiple-Choice Test Items

The *stem* of a multiple-choice test item is the part that asks the question. It may be phrased either as a question or as an incomplete sentence. *Responses* or *options* are the list of choices in a multiple-choice item from which the student chooses an answer. The incorrect options of a multiple-choice item are called *distracters* or *foils* because their purpose is to distract or foil students who don't know the correct answer from choosing it.

- ☐ **Avoid repeating words between the stem and the correct answer**, a clue that test-wise 0 will pick up.
- ☐ **Avoid using *None of the above*.** A student may correctly recognize wrong answers without knowing the right answer. So use this option only when it is important that the student know what *not* to do. If you use *None of the above*, use it in more than one question, both as a correct answer and as an incorrect option.
- ☐ **Avoid using *All of the above*.** This option requires students to read every option, penalizing those in a timed testing situation who know the material but are slow readers or for whom English is a second language. Students who recognize Option A as correct and choose it without reading further are also penalized. *All of the above* also gives full credit for incomplete understanding; some students may recognize Options A and B as correct and therefore correctly choose *All of the above* even though they don't recognize Option C as correct.

Writing Good Distracters

- ☐ **The best distracters (see Jargon Alert) help diagnose where each student went wrong in his or her thinking**, as discussed at the beginning of this chapter. Identify each mental task that students need to do to answer a question correctly, and create a distracter for the answer students would arrive at if they completed each step incorrectly.
- ☐ **Use intrinsically true or at least plausible statements.** Test-wise students recognize ridiculous statements as wrong. To see if your test has such statements, ask a friend who's never studied the subject to take the test. His or her score should be roughly what would be earned from guessing randomly on every item (25 percent for a four-option multiple-choice test).
- ☐ **Each distracter in difficult items should be attractive to at least a few students.** Unless the test item assesses an essential skill that the vast majority of students should answer correctly, if no one chooses a distracter, it's a waste of reading time. You can check this after the test is administered by counting the number of students choosing each option.
- ☐ **Use distracters that will foil test-wise students who haven't learned the concept or skill being assessed.** Create a verbal association between a distracter and the stem, for example, or make one of the distracters a relatively long option.

Writing good multiple-choice items can be difficult; test publishers write, try out, and discard many, many items for each one that ends up in a published test. Exhibit 17.2 gives some examples of multiple-choice items that follow most of the suggestions in List 17.2 and assess thinking skills as well as basic understanding. But even these examples don't completely follow all the suggestions in List 17.2. Don't expect to be able to follow every suggestion all the time, and don't expect your test questions to work perfectly the first time you use them. Analyze the test results (Chapter 24), revise the test accordingly, and within just a few cycles you'll have a really good test.

Exhibit 17.2 Multiple-Choice Questions on Assessment Concepts

(Correct answers are in *italics*.)

1. Which statement refers to measurement as opposed to evaluation?
   A. *Emily got 90% correct on the math test*.
   B. Chris's test scores have increased satisfactorily this year.
   C. Justin's score of 20 on this test indicates that his study habits are ineffective.
   D. Keesha got straight As in her history courses this year.
2. Alyssa took a test on Tuesday after a big fight with her parents Monday night. She scored a 72. Her professor let her retake the same test on Thursday when things cooled off. She scored 75. The difference in her scores may be attributed to:
   A. chance or luck.
   B. lack of discrimination.
   C. lack of validity.
   D. *measurement error*.
3. People who score high on the Meyers Musical Aptitude Scale usually score low on the Briggs Biologists Aptitude Test. People who score low on the Meyers usually score high on the Briggs. Which of the figures below *most likely* represents the correlation between the two tests?
   A. .80
   B. .00
   C. –.10
   D. *–.60*
4. Choose the *most likely* correct answer to this nonsense question, based on what you know about informed guessing on tests. A drabble will coagulate under what circumstances?
   A. Only when pics increase
   B. Only when pics change color
   C. By drawing itself into a circle
   D. *Usually when pics increase, but occasionally when pics decrease*

Jargon Alert!

Interpretive Exercises

Interpretive exercises, sometimes referred to as *integrative item sets, context-dependent* or *enhanced multiple-choice* items, or *scenario testing*, consist of a stimulus, such as a reading passage, vignette, diagram, or chart, that students haven't seen before, followed by a set of objective items on that stimulus.

Interpretive exercises

Interpretive exercises (see Jargon Alert) have three defining characteristics.

The stimulus material that students respond to must be new to the students; they must never have seen it before. This requires students to apply what they have learned to a new situation, making interpretive exercises a good way to assess application skills.

Students must read or examine the stimulus material in order to answer the objective items that follow. They should not be able to answer any of the items simply from their general understanding of what they've learned. This makes interpretive exercises an assessment of application skill rather than simple conceptual understanding.

The items must be objective, with one and only one correct answer for each item. If you ask students to write or otherwise create something in response to the stimulus, you have a performance assessment, not an objective test. Chapter 16 discusses how to create effective performance assessments.

Exhibit 17.3 is an example of an interpretive exercise.

Interpretive exercises can assess skills in generalizing, inferring, concluding, problem-solving, and analysis as well as applying knowledge and skills to new situations. Performance assessments can assess these skills as well, but interpretive exercises, like all objective item formats, are more efficient, as discussed earlier in this chapter.

Interpretive exercises are not always appropriate, however. If the stimuli are reading passages, interpretive exercises may unfairly penalize students who have achieved your learning goals but are slow readers or for whom English is a second language. And, while interpretive exercises are very good for assessing some thinking skills, they cannot assess other thinking skills such as organizing, defining problems, and creating.

The key to writing good interpretive exercises is to keep in mind their three defining characteristics. List 17.3 offers additional suggestions.

K-type items

K-type multiple choice items (see Jargon Alert) have been largely discredited (Paniagua & Swygert, 2016; Parkes & Zimmaro, 2016). One reason is that they require students to read every option and study all the possible permutations.

Exhibit 17.3 An Example of an Interpretive Exercise

(Item analysis – the concept assessed by these – is discussed in Chapter 24.)

Items 1–5 refer to the item analysis information given below. The correct options are marked with a *.

Item 1	A	B*	C	D
Top third		10		
Bottom third	1	4	3	2
Item 2	A*	B	C	D
Top third	8			2
Bottom third		7	3	
Item 3	A	B	C*	D
Top third	5		1	4
Bottom third	2		4	4
Item 4	A*	B	C	D
Top third	10			
Bottom third	9	1		

Write the item number (1, 2, 3, or 4) in the space provided.

1. **4** Which item is easiest?
2. **3** Which item shows negative (*very* bad) discrimination?
3. **2** Which item discriminates *best* between high and low scores?

For the remaining items, write the option letter (A, B, C, or D) in the space provided.

1. **B** In Item 2, which distracter is most effective?
2. **A** In Item 3, which distracter *must* be changed?

List 17.3 Tips for Writing Good Interpretive Exercises

- ☐ **Keep the size of the stimulus in proportion to the questions asked.** Having students read a full page of text in order to answer only three questions is hardly an efficient use of their time. Generally, aim to ask at least three questions about any stimulus, and ask more about longer stimuli.
- ☐ **Be on the lookout for interlocking items.** They seem to crop up more often in this format.
- ☐ **Give students realistic scenarios**.
- ☐ **Be creative!** The stimulus need not be a reading passage; it can be any of the following:
  - A chart, diagram, map, or drawing with real or hypothetical information (Exhibit 17.3)
  - A brief statement written by a scholar, researcher, or other significant individual
  - A description of a real or imaginary scenario, such as a scientific experiment or a business situation
  - For foreign language courses and programs, any of the above written in a foreign language

Jargon Alert!

K-Type Items

K-type or *complex* multiple-choice items consist of a stem followed by options, more than one of which may be correct. Those options are followed by additional options that are combinations of the initial options (*A and B*, *A and C*, *All of the above*, and so on).

This takes a lot of time, which limits the number of questions – and learning goals – that can be included in a timed testing period. The format also penalizes students who are slow readers or for whom English is a second language.

Another reason that K-type items are largely discredited is that they give test-wise students clues from the options and combinations presented. A test-wise student who recognizes that Option A is wrong, for example, will eliminate all the other options that include A. Similarly, a test-wise student who recognizes that Option B is correct will eliminate all options that don't include B.

The only legitimate reason to give students K-type items is to prepare them for a licensure or certification exam that uses this format. In this situation, your course or program's learning goals should include the analytical reasoning skills needed to tackle these questions successfully, and your students should have plenty of opportunity to learn and practice those skills before taking your tests.

Matching items

If, as you write a multiple-choice test, you find yourself writing several items with similar options, consider converting them into matching items (see Jargon Alert), which are especially efficient multiple-choice questions. Because students need to read only one set of options to answer several items, students can often answer five well-written matching items more quickly than five multiple-choice items, giving you more assessment information in the same amount of testing time. Matching items are also faster to write than multiple-choice items because you don't have to come up with a fresh set of distracters for each item.

Jargon Alert!

Matching Items

Matching items (also known as *extended matching* or *R-type* items) are a set of multiple-choice items with identical options.

Matching items are a good way to assess certain kinds of basic knowledge. Students can match terms and definitions, causes and effects, people and their achievements, foreign words and their English translations, or tools and their uses.

Matching items can also assess some thinking skills, especially the ability to apply what students have learned to new situations and the ability to analyze interrelationships. Students can match concepts with examples, causes with likely effects, symptoms with likely diagnoses, diseases with potential treatments,

and hypothetical problems with the tools, concepts, or approaches needed to solve them.

List 17.4 offers suggestions to help create good sets of matching items . . . possibly quite different from those you may have used or seen in the past. Exhibit 17.4 is an example of a matching set that follows these suggestions and assesses application skills as well as conceptual understanding.

List 17.4 Tips for Writing Good Matching Items

- ☐ **A matching set should consist of homogenous items.** Every option in the answer key should be a plausible answer for every item in the set. Otherwise, test-wise students will quickly eliminate implausible answers, while students who are less test-wise will read and consider the full set of responses over and over.
- ☐ **Allow students to use each option more than once or not at all.** A perfect match (in which each option is the answer for exactly one item) gives an unfair advantage to test-wise students, who will cross out each option as it is chosen and then guess among the options that are left. A perfect match also gives an unfair disadvantage to students who misunderstand one item but truly know all the other answers; if they choose one incorrect answer, they must, by process of elimination, choose a second incorrect answer, because it will be the only option left.
- ☐ **Make it easy for students who have learned the concept or skill to find the correct answer.** Make the longer statements the stems, and limit the answer key to single words or short phrases. For example, list definitions as the stems and the terms they define as the options. Otherwise, students will need to continually scan through a list of lengthy options to find the correct answers, which penalizes those who have learned the material but are slow readers or for whom English is a second language. Limit the number of matching items in a set to no more than 10 or so, and keep the entire exercise on one page or screen. Arrange the options in a logical order (usually alphabetically).
- ☐ **Give clear directions.** In an introductory sentence, explain how the stems and options are related (for example, ask students to match each theory with the person who conceived it). *Point out that options may be used more than once or not at all*; your students may have not seen this kind of matching set before. Consider giving the set of stems and the set of options explanatory titles (for example, *Theory* and *Author*).
- ☐ **Be inventive!** Matching sets need not be lists of words or phrases. Students can match concepts with their symbols, pictures of objects with their names, or lettered parts of a diagram, drawing (say, a microscope or cell), map, or chart with their functions, uses, or meaning.

Jargon Alert!

True-False Items

True-false or *binary* items are multiple-choice items with only two options.

True-false items

The most common use of true-false items (see Jargon Alert) is to assess basic knowledge: Is a given statement correct

Exhibit 17.4 Matching Items from a Nursing Research Methods Test

In this set of matching items, some options may be used more than once or not at all. Correct answers are in **boldface.**

Match each measurement (1–5) with its level of measurement. (A–D)

A. Interval
B. Ordinal
C. Nominal
D. Ratio

D 1. Fluid intake, in ounces, of a post-surgical patient
C 2. Religious affiliation
D 3. Medication dosage
C 4. Type of adjuvant therapy (chemotherapy, hormonal therapy, or radiation therapy)
B 5. Level of patient advocate support for a patient (very supportive, moderately supportive, somewhat supportive, not supportive)

Adapted with permission from test questions written by Dr. Christina Barrick, Associate Professor of Nursing, Towson University

or not? But they can also be used in other situations that have just two possible answers: Is a statement fact or opinion? Is this an example of direct or indirect evidence of student learning?

True-false items may be an appropriate alternative to the K-type item format, discussed earlier, when students must know that more than one option may be correct. For example, students may need to know that a symptom may be a sign of several diseases but not others. You can assess their knowledge of this by asking them a series of true-false items, each asking whether the symptom is a sign of a particular disease. But otherwise, true-false items have such serious shortcomings that they should be used rarely.

- Students who haven't learned the material have a high probability (50 percent) of guessing the correct answer.
- Unlike multiple-choice and matching items, true-false items give no diagnostic information; they give no clues about where students who answer incorrectly went wrong in their thinking.
- It's difficult to write true-false items assessing thinking skills, although they can be used in interpretive exercises.
- For classic true-false items – those giving true or false statements – it can be very difficult to write unambiguous, unqualified statements that are either always definitely true or always definitely false.
- Students may correctly recognize a false statement without knowing its true counterpart.

If true-false items appear to be appropriate for your situation, the suggestions in List 17.5 will help make the best of them.

List 17.5 Tips for Writing Good True-False Items

- ☐ **Keep them simple.** Avoid lengthy qualifiers and broad generalizations, which can be confusing and hard to make plausible as true or false statements.
- ☐ **Use them only to assess important learning goals.** It's easy for true-false items to descend into trivia.
- ☐ **Avoid negative and double-negative statements**. These are especially confusing in a true-false format.
- ☐ **Keep the proportion of true statements close to but not exactly 50 percent.** Test-wise students will scan the number of true statements they've marked and use that to decide how to guess on the items they don't know.

Jargon Alert!

Completion or Fill-in-the-blank Items

Completion items are multiple-choice items with no options. They pose questions that students answer with a word, number, or symbol. *Fill-in-the-blank* items are completion items posed as a sentence with a missing word, number, or symbol.

Completion and fill-in-the-blank items

To be truly objective items, completion items (see Jargon Alert) should have *only one correct answer*. Recall the definition of an objective test given at the beginning of this chapter: A test is objective if a reasonably competent eight-year-old armed with an answer key can score it accurately. True completion items can be scored in this fashion. Many short-answer items are really *subjective*, with a number of acceptable answers that require professional judgment to score. While such subjective items may be an appropriate part of an assessment program, they take considerable time to score and provide limited information. Performance assessments, discussed in Chapter 16, may be a better choice.

Completion items are a good choice to assess those essential facts that *must* be memorized and should not be guessed from multiple-choice items. They are also appropriate when the correct answer would be easy for students to recognize in a multiple-choice format. Completion items are widely used in mathematics, for example, when a test-wise student might deduce the correct multiple-choice answer by working backward from each option. They can be a good way to develop multiple-choice distracters for future tests: Simply choose the most frequently chosen incorrect answers as distracters.

Truly objective completion items rarely assess thinking skills except in mathematics. Because scoring is difficult to automate, this format is not a good choice for large-scale assessment programs.

List 17.6 offers suggestions for writing truly objective completion or fill-in-the-blank items.

List 17.6 Tips for Writing Good Completion or Fill-in-the-Blank Items

- ☐ Design them so one specific word, number, or symbol is the *only* correct answer.
- ☐ Keep all blanks or spaces for recording answers of uniform length. Blanks or spaces of varying length give test-wise students clues. To facilitate scoring, have students record all their answers in a column on one side of the page. If you are using fill-in-the-blank items, make the blanks in the sentences very short placeholders and have students write their answers in a column of longer blanks.
- ☐ If you are using fill-in-the-blank items, structure sentences so the blanks are toward the ends of the sentences. The sentences will be easier for students to understand than if the blanks are at the beginning.
- ☐ Avoid lifting sentences out of a textbook. Too often the resulting items are ambiguous or focus on trivia.

Pulling an objective test together

Before assembling items into a test, review them in terms of the following:

- **Do the items follow the test blueprint?** Do they each assess an important learning goal, or are any of them trick questions that ask about trivia?
- **Are the items at an appropriate reading level?** Other than vocabulary terms that you are specifically assessing, are the items simple, clear, and straightforward? Are they free of excessive verbiage?
- **Would experts agree on the answers?**
- **Do the items appear to be of appropriate difficulty?**
- **Are there any interlocking items** or items with any other clues for test-wise students?

Next, order your items. The first items should be the easiest ones, to reassure test-anxious students, and quickly answered, to help those who aren't test-wise. The last items should be the most difficult and the most complex (requiring the most thinking time). Interpretive exercises often go toward the end.

Write directions that provide the information in List 17.7. Finally, let the test sit for 24 hours, and then proofread it one last time. Prepare the scoring key *before* the test is duplicated or posted online, as the process of preparing the key can identify typos and unclear items missed in earlier readings.

List 17.7 Information to Include in Test Directions

- ☐ The purpose of the test
- ☐ How the answers will be scored
- ☐ How to answer (Can they choose more than one answer? Should they choose the one *best* answer? Is guessing encouraged?)
- ☐ How to record answers
- ☐ Any time limits (If the test is lengthy and timed – a two-hour final exam, for example – you may want to suggest time limits for each section.)

Should students explain their answers to objective items?

As noted earlier, a major concern with objective items, especially true-false items, is the possibility that students who haven't learned the material can still guess the correct answer. One way to solve this problem is to ask students to write brief explanations of why they chose their answer. For true-false items, students can be asked to correct any statements they mark as false.

While this does eliminate the possibility of students guessing the correct answer, it also removes one of the fundamental advantages of objective items: Their efficiency. Students won't be able to answer as many questions in a given amount of testing time, so the test will assess fewer learning goals. And, rather than having a scanner or a competent eight-year-old score the tests, you must read every answer and use your professional judgment to decide which are correct and which are not, which takes far more of your time.

If you want students to explain their answers, consider instead giving them an assignment or essay question for which they must compose a more complete written response. This will elicit deeper thought, give you richer assessment information, and give your students a better learning experience. See Chapter 16 for suggestions.

Should students be encouraged to guess?

In a word, yes. Test-wise students readily guess on items of which they're unsure. They know that, if they can eliminate even one option as implausible, they raise their odds of guessing correctly beyond random chance. If students who aren't test-wise aren't encouraged to guess, they're being unfairly penalized for not being sufficiently test-savvy.

Time to think, discuss, and practice

Write each of the following for a unit or concept that you teach or have studied, following as many of the guidelines in this chapter as you can. Share your drafts with group members for feedback and suggestions.

1. Six multiple-choice items
2. A set of matching items
3. An interpretive exercise