Chapter

1

Standardized Tests

THERE IS NO SUCH THING AS A "POOR STANDARDIZED TEST-TAKER"

In our opinion, anyone who has ever uttered the sentence, "I'm not a good standardized test-taker" should have said, "I don't prepare well for standardized tests." Our hope is that after reading this book you will have no need to say either.

We do not believe there is any such thing as a "poor standardized test-taker." The phrase itself creates the impression of a mysterious innate quality (or lack thereof) that renders someone constitutionally incapable of taking a standardized test. In our experience, students often offer this as an explanation for a low score, and they do so in a manner that is almost matter-of-fact and fatalistic as though that's just it, game over, they're bad at taking standardized tests, and they always will be because it's . . . part . . . of . . . who . . . they . . . *are*!

Certainly, standardized tests pose different challenges than other types of tests you have taken. However, there is no fundamental skill required for taking and doing well on a standardized test that you cannot master. It's all about proper preparation, and that's something you can do, with some guidance.

It's likely you've already taken a standardized test, such as the SAT or ACT, and you've probably

talked with others about standardized tests. So, the first of many questions we will pose is this: What makes a standardized test a standardized test?

Do you know? Do you have an idea or a guess?

Don't read any further until you stop and think about your answer!

Do you have an answer yet? If yes, continue reading to see how your answer compares to ours. If not, reflect for a few moments more before proceeding.

Critical Comment: Active reflection is a valuable skill to develop. You should train yourself to periodically stop reading and think about the material. Ask and answer your own questions. How is this material related to what I already know? What is its relevance? This is a learning method called elaboration, and it is associated with deeper learning and better long-term memory retention (for more information, refer to Chapter 7).

Our answer: A standardized test differs from other types in its purpose and design. The ultimate purpose of a standardized test is, as the name implies, standardization; it provides a standard for comparison. Standardized tests are designed to evaluate and then compare the aptitudes or competencies of a diverse population of individuals (e.g., students from different institutions who have different educational backgrounds). Therefore, a standardized test must be: (1) representative of a domain of knowledge, (2) dependable with regard to the format and scoring, and (3) consistent in terms of testing conditions (*Standardized Assessment: A Primer*, 2011).

Critical Comment: If ever a word is unfamiliar to you, no matter what you're reading, be sure to look up the definition in a dictionary. Vocabulary is often tested on standardized exams. Don't memorize words and hope that you will recognize them on the test. Understand what you are studying—give it meaning! Incorporate new vocabulary into your everyday conversations. Make it fun. Stump your friends. Subscribe to a word-a-day service.

So, the question is: What is it about these qualities of standardized tests that makes them more challenging than the majority of exams and quizzes you have experienced thus far? To appreciate the answer to that question, you need a basic understanding of several aspects of test design.

TEACHER-GENERATED TESTS VERSUS STANDARDIZED TESTS

Teacher-generated tests are generally designed to evaluate your learning of the course material in order to assign a course grade. They are usually written for a specific course (e.g., BS101) within a specific discipline (e.g., bioscience) and are typically administered in the same classroom as the course is taught under the supervision of the faculty member responsible for teaching that course. The focus of teacher-generated exams is limited in content and the questions tend to emphasize knowledge of facts, specific details or specific concepts. They are generally not intended to measure learning across different courses or disciplines. In other words, if you take an exam in course SOS202 (social science), you are not likely to be tested on content you learned in a previous bioscience course (BS101) or history, chemistry, or algebra course unless that content was specifically covered in SOS202.

For teacher-generated course exams, the similarities or continuities between the learning environment and the testing environment provide numerous conscious and subconscious clues and memory cues (triggers) to help you recall information on the exam. For example, you might remember the sound of the professor's voice as she was explaining the concept you need to understand to answer a test question. Or you might visualize the professor standing in the front of the room explaining his PowerPoint slide to the class, which then triggers your recall of the correct answer to the question.
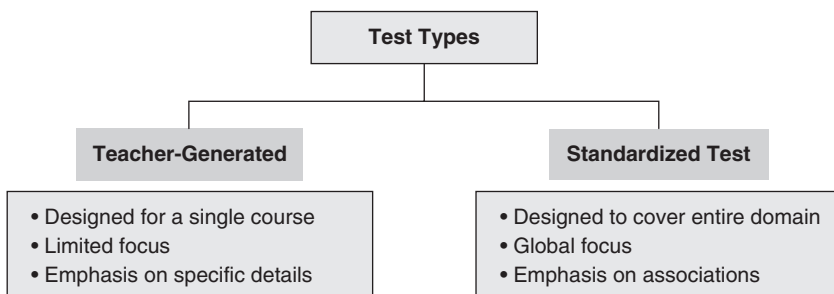
In contrast, standardized tests are designed to measure knowledge across courses, subjects, and disciplines; therefore, they generally cover a much broader domain of information. Their focus tends to be more global with an emphasis on associations, integration, and application (see Figure 1-1). They may even be intended to evaluate specific cognitive skills, such as problem solving, verbal reasoning, or analytical thinking. Moreover, standardized test questions are written and edited by a group of content experts and professional editors—people you have never met—using established test construction techniques developed with the intent of reducing your ability to use cues and clues to select the correct answers.

Finally, the questions themselves usually differ between the two types of exams. Test questions can be categorized by the type of cognitive task
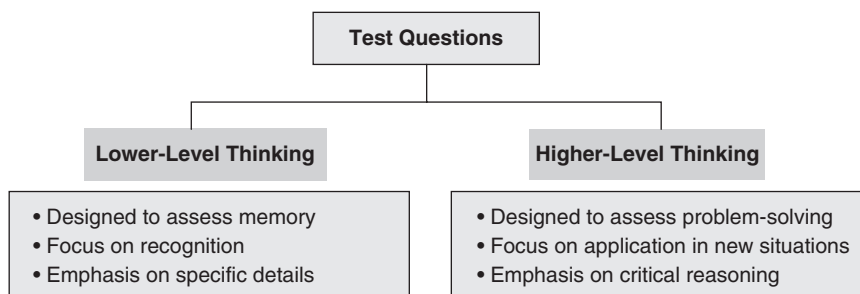
Figure 1-1

The Two Primary Types of Tests

```
                          ┌─────────────┐
                          │  Test Types │
                          └─────────────┘
                  ┌───────────────┴───────────────┐
          ┌──────────────────┐            ┌──────────────────┐
          │ Teacher-Generated │            │ Standardized Test │
          └──────────────────┘            └──────────────────┘
```

Teacher-Generated	Standardized Test
• Designed for a single course • Limited focus • Emphasis on specific details	• Designed to cover entire domain • Global focus • Emphasis on associations

required to answer them. In the simplest sense, they are divided into two groups: low-level items or high-level items. Coursework frequently focuses on new terms, facts, and concepts. To assess a student's ability to recognize a term or recall a fact, teacher-generated tests generally ask questions that are limited in scope, depth, and complexity. Such questions are referred to as low-level thinking items. High-level thinking items require a deeper understanding of principles and concepts; they test the examinee's ability to solve a problem, evaluate a situation, make a judgment, or synthesize new ideas. In contrast to teacher-generated tests, standardized tests tend to be composed of a much higher percentage of high-level questions (see Figure 1-2). This tends to make standardized tests more challenging, especially if you're not expecting this difference.

Figure 1-2

The Two Basic Types of Test Questions

Lower-Level Thinking	Higher-Level Thinking
• Designed to assess memory • Focus on recognition • Emphasis on specific details	• Designed to assess problem-solving • Focus on application in new situations • Emphasis on critical reasoning

Many students learn to prepare for and perform on teacher-generated exams using just a few basic skills:

- Read and memorize key words or phrases from the textbook.
- Memorize class handouts and notes.
- Listen for clues from the teacher about what will likely be on the test.
- Memorize old test questions.
- Answer multiple-choice questions by recognizing memorized facts and concepts.
- Answer multiple-choice questions by recognizing clues provided by the teacher.

The bottom line is that through years of conditioning you have learned how to pass teacher-dependent, teacher-created tests. Standardized tests are teacher-*independent* tests. Excelling on standardized tests requires that you develop new skills and refine old skills. In short, you need to become a test-wise person with excellent test-taking skills. Before we follow that train of thought, we want to be sure you are familiar with a few more terms and concepts related to tests—validity, reliability, and scoring.

TEST VALIDITY AND RELIABILITY

A "good" test should possess two characteristics—validity and reliability. To be valid, the test must measure what it intends to measure. To be reliable, a test must produce essentially the same results when administered at multiple sites over an extended period of time; in other words, the results must be reproducible. A test that produces statistically different scores depending on the testing location is not reliable. A test cannot be valid if it is not also reliable.

There are several types of validity, but we will mention only three: face, content, and concurrent. Test-takers often have an opinion about tests; their perception of an examination is known as face validity. Content experts design tests; their expertise ensures the exam tests the material it is supposed to test (content validity). Finally, a test should correlate well with a previously validated test (measurement instrument) that was designed to measure the same thing; this is called concurrent validity.

> **Critical Comment:**   If you're beginning to lose patience and are sitting there thinking, "Just tell me what I need to know already, so I can do well on this test!" then whether you realize it yet or not, you need this book, so keep reading!

HOW ARE STANDARDIZED TESTS SCORED?

Initial scoring of an examination produces a "raw score," which is simply the number of correctly answered questions. The *percent* correct score is obtained by dividing the number of correct responses by the number of total items on the examination. For example, if there were 150 items on the exam and you answered 114 correctly, your raw score would be 114 and your percent correct score would be 76% (114/150).

Is the interpretation of standardized test scores so simple? In a word, *No*! Two practices make the interpretation of standardized test scores more challenging—the inclusion of pilot items and normalization of scores.

Pilot Items

Pilot items are test questions that do not count in the determination of your final score. Standardized test questions go through an item validation process prior to their use in determining an actual test score (Shrock & Coscarelli, 2000). This is to eliminate invalid and/or unreliable questions. To validate a question, it is included on an actual examination to see how it performs in a real-world testing situation—how many examinees select the correct answer, how many examinees select incorrect responses and which ones, did examinees from different areas of the country perform the same, and so on. This process, known as psychometric validation, uses a lot of statistics. The good news for you is that you don't need to understand how it's done; just being aware that it occurs is enough.

HOW DO PILOT ITEMS IMPACT YOUR FINAL REPORTED SCORE?

Your reported score (i.e., the final score you receive from the testing agency) reflects the number of correctly answered *scored* items, not

inclusive of the pilot items. To better understand what this means, let's look at a specific example:

For an examination with 150 total items, 30 of which are pilot items, your reported score will be based on the number of questions you answered correctly from the 120 *scored* items, not the entire pool of 150 items. So let's say you answered 90 items correctly out of the total number of 150; what will be your reported score? That depends. It depends on how many of the questions that you answered correctly were scored items because your reported score is based only on the number correct of the 120 *scored* items. If, of your 90 correct answers, only 60 of them were among the 120 scored items, your final reported raw score would be 60 and your reported *percent* correct would be 50% (60/120) not 60% (90/150). What if all 90 questions that you answered correctly were scored items? Then, your reported raw score would be 90 and your reported percent score would be 75% (90/120).

So you can see that the number of pilot items on an exam can significantly affect your final reported score. Pilot items can also impact your *perception* of how you're doing on an exam *while* you are taking it. For example, during a standardized test, you might come across questions that seem poorly worded, extra confusing, or otherwise unexpected. It is certainly possible these represent pilot items. In order to avoid becoming unduly overwhelmed, worried, and anxious or distracted by these questions, we suggest you assume they *are* pilot items and will not be scored. Move on to the next question with no worries! We will revisit this idea in Section IV.

Executing Your Game Plan (Test Day)

To add another layer of complexity, on most standardized tests, your final reported score is not a percent correct score; it is a *normalized score*. Generally, the next question we are asked is, "What is a normalized score?"

Normalized Test Scores

To allow test scores to be interpreted in more meaningful ways, test results are converted into data that allow the relative performance of examinees to be compared. This data is referred to as normative data

and is generally described as either norm referenced or criterion refer-
enced. Although we will first discuss norm referencing, you should be
aware that many professional and licensing examinations have shifted
toward a criterion-referenced framework to assess examinees' degree of
competency (Shrock & Coscarelli, 2000).

One way to normalize a score is to take an examinee's raw score
(the number of correct responses on the *scored* items) and compare
it to the performance of the entire group of examinees that took the
same (or psychometrically equivalent) test. This comparison results in
a percentile rank. Percentile ranks range from a low of 1 to a high of 99.
Percentile ranks provide more insight into the examinee's ability than
the raw score alone. As an example, a percentile rank of 90 indicates
that the examinee's score was higher than 90% of the scores obtained
by the group. Earlier, we stated that answering 60 items correctly out
of 120 scored items results in a percent correct of 50%. If fewer than
one-fifth of all examinees scored above 50%, however, then the score
of 50% would result in a percentile score of at least 80 or greater, which
makes that score of 50% sound more appealing, doesn't it?

There's a problem with comparing examinees using the percentile
ranking system; the percentile range will always be 1–99. What if ev-
eryone who took the test answered fewer than half the questions cor-
rectly? In a percentile ranking system, no matter how low the raw and
percent scores are overall, the person who received the highest score
is always going to have the highest percentile score and that will be at
the 99th percentile. Because a percentile rank indicates only relative
performance, we would know nothing about whether any or all of
the test-takers possess a baseline (minimal) level of competency that
would be required to perform an important activity, such as practicing
medicine or nursing.

Criterion-referenced scoring offers an advantage to individuals who
use test results to make judgments about competency. Instead of com-
paring examinees to one another, criterion-referenced scoring com-
pares each examinee score to a predetermined standard, which defines a
"mastery" level of knowledge. The "cut score" (i.e., the score that prede-
termines pass/fail) is established prior to administration of the test by a
panel of content experts using—you guessed it—a series of psychomet-
ric exercises and statistical standards to ensure that the cut score is valid.
If an examinee's score is above the cut score, the examinee passes. Using

criterion-referenced scoring, it is possible that all examinees could pass (or fail) the examination. The advantage to the examinee of criterion-referenced testing is that the score they obtain, and need to achieve to pass the exam, is determined by comparing it to a predetermined level of proficiency rather than just other test-takers.

SUMMARY

This chapter described the basic elements of standardized tests and contrasted them to teacher-generated tests. Standardized exams test a larger domain of knowledge and breadth of skills; they are designed based on predetermined formats, delivered under specific conditions, and scored using psychometric standards.

We have presented this information for the simple reason that individuals struggle with standardized tests because (1) they have trained themselves to pass teacher-generated, not standardized, tests and (2) they don't realize that studying for standardized tests requires different study methods.

Before proceeding to Chapter 2, please reflect on the following questions:

- What was the primary take-away message for you from this chapter?
- What information will be the easiest for you to include in your exam preparation?

Be sure to complete the Chapter 1 Activity before starting to read Chapter 2.

CHAPTER 1 ACTIVITY

Becoming a Better Standardized Test-Taker

In Chapter 1—Standardized Tests—we stated that "there is no such thing as a poor standardized test-taker"; however, there are individuals that could prepare better for a standardized test. Use the following questions as a guideline and comment on how you will approach each of the situations listed to help you become a better test preparer.

Table	
1–1	**Activity 1**

a. When reading new material, what can I do to give it more meaning for me?

b. What characteristics differentiate teacher-generated from standardized tests?

c. What makes higher-level thinking questions more challenging?

d. What techniques do I use to prepare for teacher-generated tests?

e. What techniques will I use to prepare for my upcoming standardized test?