

THE BASICS

“**A**nd in the beginning, there was . . . a test.” Not really, but, in a little bit, you will be surprised to learn just how long ago this whole measurement thing started. That’s what the first part of *Tests & Measurement for People Who (Think They) Hate Tests & Measurement* is all about—a little history, an introduction to what kinds of tests there are and what they are used for, and then something about how to use this book.

You’re probably new to this, and we’re sure you couldn’t wait for this course to begin. ☺ Well, it’s here now, and believe it or not, there’s a lot to learn that can be instructive and even fun—and immeasurably valuable. Let’s get to it.

Do not copy, post, or distribute

Do not copy, post, or distribute



WHY MEASUREMENT?

An Introduction

Difficulty Index 😊😊😊😊😊
(the easiest chapter in the book)

LEARNING OBJECTIVES

After reading this chapter, you should be able to

- List the important events and milestones in the history of tests and measurement.
- Describe the role of testing and social science measurement in our lives and in academia.
- Identify important caveats and issues when studying tests and measurement.
- Give the steps in the developmental process for a good test.
- Defend the need for a stand-alone course in tests and measurement.
- Explain how to use this book and what all those cute smiley faces mean.

It's been happening to you, and you've been doing it, since you were very young. Your whole life, you've been tested.

And it started even before you were born. While she was pregnant, your mom probably had the doctor assess you using ultrasound technology to measure your organs and evaluate whether your development was “normal.” (From your very first test-taking experience, measurement experts have insisted on comparing you to other people!) When you were born, the doctor administered the APGAR to assess your *Appearance* (or color), *Pulse* (or heart rate), *Grimace* (or response to stimulation), *Activity* (or muscle tone), and *Respiration* (or breathing). You were also screened (and it’s the law in almost every state) for certain types of metabolic disorders (such as PKU, or phenylketonuria).

Then there may have been personality tests (see Chapter 10), spelling tests (see Chapter 12), statewide tests of educational progress (see Chapter 7), the ACT (American College Test) or the SAT (which actually is not an acronym—see Chapter 8 for more on this), and maybe even the GRE (Graduate Record Exam). Along the way, you might have received some career counseling using the SVIB (Strong Vocational Interest Blank) and perhaps a personality test or two, such as the MMPI (Minnesota Multiphasic Personality Inventory) or the MBTI (Myers-Briggs Type Inventory).

My, that’s a lot of testing, and you’re nowhere near done.

You’ve still probably got a test or two to complete once you graduate from school, perhaps as part of a job application, for additional studies, or for screening for a highly sensitive job as a secret agent. And we haven’t even begun to list all the medical tests you have and will have been subject to in your lifetime!

Testing is ubiquitous in our society, and you can’t pick up a copy of the *New York Times*, *Chicago Tribune*, or *Los Angeles Times* without finding an article about testing and some associated controversy.

The purpose of *Tests & Measurement for People Who (Think They) Hate Tests & Measurement* is to provide an overview of the many different facets of testing, including a definition of what *tests and measurement* is as a discipline and why it is important to study, the design of tests; the use of tests, and some of the basic social, political, and legal issues that the process of testing involves. And when we use the word *test*, we are referring to any type of assessment tool, assessing a multitude of behaviors or outcomes.

This first part of *Tests & Measurement for People Who (Think They) Hate Tests & Measurement* will familiarize you with a basic history of testing and what the major topics are that we as teachers, nurses, social workers, psychologists, parents, and human resource managers need to understand to best negotiate our way through the maze of assessment that is a personal and professional part of our lives.

Let’s start at the beginning and take a brief look at what we know about the practice of testing and how we got to where we are.

A FIVE-MINUTE HISTORY OF TESTING

First, you can follow all this history stuff by using the cool time line for what happened when, beginning at the bottom of this page and appearing throughout the chapter. Here's a summary.

Imagine this. It's about 2200 years BCE (Before the Common Era), and you're a young citizen living in a large city in China looking for work. You get up, have some breakfast, walk over to the local "testing bureau," and sit down and take a test for what we now know as a civil service position (such as a mail carrier). And at that time, you had to be proficient in such things as writing, arithmetic, horsemanship, and even archery to be considered for such a position. Must have been an interesting mail route.

Yep—testing in one form or another started that long ago, and for almost 3,000 years in China, this open (anyone could participate), competitive (only the best got the job) system proved to be the model for later systems of evaluating and placing individuals (such as the American and British civil service systems that started around 1889 and 1830, respectively).

Interestingly, this system of selection was abandoned in China around the turn of the 20th century, but we know from our own experience that the use of testing for all different purposes has grown rapidly.

Testing is on the increase by leaps and bounds and it's not getting any cheaper. The Brookings Institution think tank has estimated that states spend \$1.7 billion yearly on just those federally mandated state tests for the K–12 crowd. That's a ton of money, and the entire endeavor is expected to get even more expensive as tests and their use in school accountability continue to grow in popularity among our leaders.

In terms of important doings related to tests, not much of a formal or recorded nature occurred before the middle of the 19th century, but by about the end of the 19th century, along comes our friend Charles Darwin, whom you may know from some of your other classes as the author of *On the Origin of Species* (available in the first edition, very good shape, for only about \$440,000; but the shipping is free). This book (of which only 11 copies of the first edition have survived) is a groundbreaking work that stressed the importance of what Darwin called "descent with modification" (which we now call evolution). His thesis was that through the process of variation, certain traits and attributes are selected (that is, they survive while others die out), and these traits or attributes are passed on from generation to generation as organisms adapt.

So why are we talking about Charles Darwin and biology in a tests and measurement book? Two reasons.

First, Darwin's work led to an increased interest in and emphasis on individual differences—and that's what most tests examine. And second, Darwin's cousin (how's that for a transition?) Francis Galton was the first person to devise a set of tools for assessing individual differences in his *anthropometric* (measurements of the human body) lab where one could have all kinds of variables measured, such as height, weight, strength, and even how steady you can hold your hands. His motto was, "Wherever you can, count." (And, by the way, Sherlock Holmes's motto was "Data! Data! Data!" They must have been *very* busy guys.)

Once physical measurements were being made regularly, it was not long before such noted psychologists as James Cattell were working on the first "mental test." Cattell was a founder of the Psychological Corporation in the early 1920s, now known as one of the leading publishers of tests throughout the world.

When we get to the 20th century, testing and measurement activity really picks up. There was a huge increase in interest devoted to mental tests, which became known as intelligence tests (and, less accurately, IQ tests) and also included the testing of cognitive abilities such as memory and comprehension. More about this in Chapter 9.

A major event in the history of testing occurred around 1905, when Alfred Binet (who was then the minister of public instruction in Paris) started applying some of these new tools to the assessment of Parisian schoolchildren who were not performing as well as expected. Along with his partner, Theodore Simon, Binet used tests of intelligence in a variety of settings—and for different purposes—beyond just evaluating schoolchildren's abilities. Their work came to the United States in about 1916 and was extended by Lewis Terman at Stanford University, which is why this still popular intelligence test is called the Stanford-Binet. (As we write these words, by the way, Bruce is listening to Alvin and the Chipmunks, those adorable hit record makers from the 1950s and 1960s and animated movie stars from the 2000s. It occurs to us that the three chipmunks are named Alvin [kind of like Alfred], Theodore, and Simon. Wonder if their creator, Ross Bagdasarian, was a historian of intelligence tests.)

As always, necessity is the mother and father of invention, and come World War II, there was a huge increase in the need to test and classify accurately those thousands of (primarily) men who were to join the armed services. (This occurred around World War I as well but with nowhere near the same amount of scientific deliberation.) *Intelligence* was one of the primary traits of interest on these tests and a strong correlation between scores on the military's intelligence test and the eventual rank of test takers was early evidence used to argue for the accuracy of intelligence tests.

And, as always, intense efforts at development within the government usually spill over to civilian life, and after World War II, hundreds of different types of tests

were available for use in the civilian sector and made their way into hospitals, schools, and businesses. Indeed, we have come a long way from spelling tests.

While all these mental and ability tests were being developed, increased attention was also being paid to other dimensions of psychological functioning, such as personality development (see Chapter 10). People might be smart (or not smart), but psychologists also wanted to know how well adjusted they were and whether they were emotionally mature enough to assume certain important responsibilities. Hence, the field of personality testing got started in earnest (around World War I) and certainly is now a major component of the whole field of tests and measurement.

But our brief history of testing does not stop with intelligence or personality testing. As education became more important, so did evaluating achievement (see Chapter 7). For example, in 1937, the then-called Stanford Achievement Tests (or SATs) became required for admission to Ivy League schools (places such as Brown, Yale, and Princeton)—with more than 2,000 high school seniors taking the exam. Another example? In 1948, the Educational Testing Service (known as ETS) opened, almost solely to emphasize the assessment of areas other than intelligence. These are the folks that bring you today's SAT, GRE, and the always popular and lovable Test of English as a Foreign Language (or TOEFL)—all taken by millions of students each year.

It's no wonder that services offering (and sometimes guaranteeing) testing success began to proliferate around 1945 with Stanley Kaplan. A very smart New Yorker (who was denied admission to medical school), Kaplan started tutoring students in the basement of his home for \$0.25 per hour. His success (and it's still a hotly debated issue whether raising test scores after instruction really indicates an increase in actual ability or knowledge) led him to create an empire of test centers (sold off for a bunch of millions to a big test company) that is still successful today.

During the COVID-19 pandemic, it became difficult to take college admissions tests like the SAT, ACT, and GRE. As a result, many college programs and whole universities stopped requiring these tests when applying for admission. This changed the way that admissions officers and faculty thought about criteria for admission and predictors of success in college. Many had, for years, questioned the usefulness and fairness of these tests to decide who gets into college, especially the top programs, so there have always been reasons beyond the barriers imposed by COVID-19 to not place a lot of weight on standardized test scores. The question is whether these colleges will ever return to requiring test scores for admission decisions. It looks as if many will not and the role of these tests may have been permanently reduced. What will be the effect of that?

The most common assessment that goes on in education, though, is not through standardized tests to predict college performance or demonstrate school effectiveness, it is the use of good old-fashioned classroom tests made by classroom teachers

for their own students. It turns out that there are research-based best practices for making a good test and Chapters 12 and 13 talk about how to write a good quiz or create a scoring rubric that will help to make the grades students get for giving a speech in class less subjective.

Today, thousands and thousands of tests (and hundreds of test publishers—see Appendix A) measure everything from Advanced Placement Examination in Studio Art, which is designed to measure college-level achievements in studio arts, to the Health Problems Checklist, which is used to assess the health status and potential health problems of clients in psychotherapy settings.

And a new emphasis on the study of neuroscience has led to new evaluative efforts that explore and assess the impact of brain behavior on performance and enable an intense look at the role and function of testing—not without a great deal of controversy about topics such as online testing, fair testing using a common core as the basis for educational valuation, high-stakes testing, test bias, and more.

SO, WHY TESTS AND MEASUREMENT?

This question has a pretty simple answer, but simple does not mean lacking in complexity or significant implications.

No matter what profession we enter, be it teaching, social work, nursing, or any one of thousands more, we are required to make judgments every day, every hour, and in some cases, every few minutes about our work. We do it so often that it becomes second nature. We even do it automatically.

In the most straightforward of terms, we use a test (be it formal or informal) to measure an outcome and make sense of that judgment. And because we are smart, we want to be able to communicate that information to others. So if we find that Deion got 100% on a spelling test or a 34 on his ACTs, we want everyone who looks at that score to know exactly what it means.

For example, consider the teacher who records a child's poor grade in math and sends home some remedial work that same evening, the nurse who sees a patient shivering and takes their temperature, or the licensed clinical social worker who recognizes a client has significant difficulties concentrating and administers a test to evaluate their ability to stay on task and designs an intervention based on that evaluation. These people all recognize a symptom of something that has to be looked into further, and they take appropriate action.

What all these professionals have in common is that in order for them to take action to help the people with whom they work, they need to first assess a particular behavior or set of behaviors. And to make that assessment, they use some kind of test (such as a standardized test in the case of the nurse or home-made

test, as in the teacher's case) to gather information. Then, based on their training and experience, they use that information to make a decision as to what course of action to take.

For our purposes here, we are going to define a **test** as a (pick any of the following) tool, procedure, device, examination, investigation, assessment, or measure of an outcome (which is usually some kind of behavior, even if the “behavior” is getting a certain score). A test can take the form of a 50-question multiple-choice history exam or a 30-minute interview of parents on their relationships with their children. It can be a set of tasks that examine how good someone is at fitting together blocks into particular designs or an attitude survey about whether they prefer multigrain Cheerios® to plain Cheerios®. (The right answer should be *plain*.) We use tests that come in many different forms to measure many different things.

What We Test

We test many, many different things, and the thousands of tests that are available today cover a wide range of areas. Here's a quick review of some of the content areas that tests cover. We'll go into greater detail on each of these in Part II of *Tests & Measurement for People Who (Think They) Hate Tests & Measurement*.

We'll define these different general areas here, and in Table 1.1 you can see a summary along with some real-world examples.

Achievement tests (covered in Chapter 7) assess an individual's level of knowledge in a particular domain, like in school. For example, your midterm in history was an achievement test. As Bruce's grandfather would say, achievement tests measure “book learnin'.”

Personality tests (covered in Chapter 10) assess an individual's unique and stable set of characteristics, traits, or attitudes. You may have taken an inventory that determined your level of introversion or extraversion. In naming Chapter 9, we use the term *psychological* tests, which is broader and allows us to include other measures that aren't technically about personality.

Aptitude tests (covered in Chapter 8) measure an individual's potential to succeed in an activity requiring a particular skill or set of skills. For example, you may take an aptitude test that assesses your potential for being a successful salesperson. Aptitude tests predict the future.

Ability or intelligence tests (covered in Chapter 9) assess one's level of skill or competence in a wide variety of areas. For example, intelligence tests are viewed as measures of ability (but don't be fooled by the name of a test—there are plenty of intelligence tests that are also seen as being aptitude tests—see the upcoming box!).

Neuropsychological tests (covered in Chapter 10) assess the functioning of the brain as it relates to everyday behaviors, including emotions and thinking.

TABLE 1.1 An Overview of What We Test and Some Examples of Such Tests

Type of Test	What It Measures	Some Examples
Achievement	Level of knowledge in a particular domain	<ul style="list-style-type: none"> • Closed High School Placement Test • Early Childhood Assessment • Norris Educational Achievement Test • Test of Adult Basic Education
Personality	Unique and stable set of characteristics, traits, or attitudes	<ul style="list-style-type: none"> • Achievement Motivation Profile • Aggression Questionnaire • Basic Living Skills Scale • Dissociative Features Profile • Inventory of Positive Thinking Traits
Aptitude	Potential to succeed	<ul style="list-style-type: none"> • Differential aptitude tests • Scholastic Aptitude Scale • Aptitude Interest Category • Evaluation Aptitude Test • Wilson Driver Selection Test
Ability or intelligence	Skill or competence	<ul style="list-style-type: none"> • Wechsler Intelligence Scale for Children • Stanford-Binet Intelligence Scales • Cognitive Abilities Test • General clerical ability tests • School Readiness Test
Neuropsychological	How your brain works	<ul style="list-style-type: none"> • Boston Naming Test • Cognitive Symptoms Checklist • d2 Test of Attention • Kaplan Baycrest Neurocognitive Assessment • Ruff Figural Fluency Test
Vocational or career	Job-related interests	<ul style="list-style-type: none"> • Adaptive Functioning Index • Career Interest Inventory • Prevocational Assessment Screen • Rothwell Miller Interest Blank • Vocational Adaptation Rating Scales

Note: You can find out more about many of these tests by going to the Buros Center for Testing at buros.org.

Finally, **vocational or career tests** (covered in Chapter 11) assess an individual’s interests and help classify those interests as they relate to particular jobs and careers. For example, you may have taken a vocational test that evaluates your level of interest in the health care professions or the culinary arts (or both, which is maybe how we got the name Dr. Pepper?).

There is always a great deal of overlap in the way people categorize particular types of tests and what they assess. For example, some people consider intelligence to be an ability (and would place it under ability tests), whereas others think of it as an achievement test because one aspect of intelligence is the ability to learn or retain information. Or aptitude tests can end up as ability tests as well as personality tests, or they can stand all on their own. And think of college admissions tests, like the ACT or SAT. They ask questions that measure learned information (which sounds like achievement), but they are used to predict success in college (which sounds like aptitude).

So what's right? They are all right. The way we classify tests is strictly a matter of organization and convenience and even a matter of how they are used. The definitions and examples given here reflect the current thinking about tests and measurement. Others feel differently. Welcome to the real world.

Why We Test

Now you know that there are different forms of tests and that there are many different areas of human performance and behavior that are tested regularly. But for what purpose? Here's a summary of the five main reasons we measure people (and there are surely more).

Tests are used for *selection*. Not everyone can be a jet pilot, so only those humans (and some smart monkeys) who score at a certain level of performance on physical and psychological assessments will be selected for training.

Tests are used for *placement*. Upon entering college, not everyone should be in the most advanced math class or in the most basic. A placement test can determine where the individual belongs.

Tests are used for *diagnosis*. An adult might seek out psychological counseling, and the psychologist may administer a test or group of tests that helps diagnose any one of many different mental disorders. Diagnostic tests are also used to identify individual strengths.

Tests are used to *classify*. Want to know what profession might suit you best? One of several different tests can provide you with an idea of your aptitude (or future potential) for a career in the culinary arts, auto mechanics, medicine, or child care.

Finally, tests are used in *research*. To find relationships among variables in the social sciences, tests are often used to assign scores—meaningful quantities—to those variables and then statistical analyses are conducted on those scores.

So, what are tests used for? Tests are used widely for a variety of purposes, among them *selection, placement, diagnosis, classification, and for measuring research variables*.

What Makes a Good Test Good?

Regardless of the purpose of a test, there are two characteristics of quality. A good test is *valid* and a good test is *reliable*. We will talk about these concepts throughout this book, but, for now, think of validity as the characteristic of a test that measures what it is supposed to; it works as intended. Reliability refers to a test that produces a score that does not vary randomly; it produces scores that represent typical performance for each person who takes it.

SOME IMPORTANT REMINDERS

You'll learn many different things throughout *Tests & Measurement for People Who (Think They) Hate Tests & Measurement* (at least we sure hope you will). And with any vibrant and changing discipline, there are always discussions both pro and con about different aspects of the subject. But there are some constants as well, as presented in the following:

- *Some behaviors can be observed more closely and more precisely than others.* It's pretty easy to measure one's ability to add single digits (such as $6 + 5 = ?$), but to understand *how* one solves (not *if* one can solve) a simple equation is a different story. The less obvious behaviors take a bit more ingenuity to measure, but that's part of the challenge (and delight) of doing this.
- *Our understanding of behavior is only as good as the tools we use to measure it.* There are all kinds of ways we try to measure outcomes, and sometimes we use the very best instruments available—and at other times, we may just use what's convenient. The development and use of the best tools takes more time, work, and money, but it gives us more accurate and reliable results. Anything short of the best forces us to compromise, and what you see may, indeed, not be what you get.
- *Tests and measurement tools can take many different forms.* A test can be paper and pencil, computer administered, self-report, observation, performance based, and so on, but often different forms give us very similar information on some outcome in which we are interested. And often, the format a test uses is determined by what it is measuring. For example, most classroom achievement tests are paper and pencil, and most tests that look at performance of motor skills are performance based. The lesson here is to select the form of test that best fits the question you are asking.
- *The results of any test should always be interpreted within the context in which it was collected.* In many communities, selected middle school students take a practice SAT test. Although some of these students do very, very

well, others perform far below what you would expect a high school junior or senior to do; this makes sense because these younger children simply have not yet had the chance to learn the material. To interpret the results of the younger children using the same metric and scoring standards as for the older children would surely not do either group any justice. The point is to keep test scores in perspective—and of course, to understand them within the initial purpose for the testing.

- *Test results often can be misused.* It doesn't take a rocket scientist to know that there have been significant controversies over how tests are used. You'll learn more about this in Part V of *Tests & Measurement for People Who (Think They) Hate Tests & Measurement*. Did you know, for example, that many non-English-speaking immigrants who tried to get sanctuary in the United States were turned away in the 1930s as having low intellectual abilities based on test scores. Tests written in English! To use tests fairly and effectively, you need to know the purpose of the test, the quality of the test, how it is administered and used, and how the results are interpreted. We'll explore these issues in *Tests & Measurement for People Who (Think They) Hate Tests & Measurement*.

Remember, no matter how interesting your theory or approach to a problem, what you learn about behavior is only as accurate and worthwhile as the integrity and usefulness of the tools you use to measure that behavior.

HOW TESTS ARE CREATED

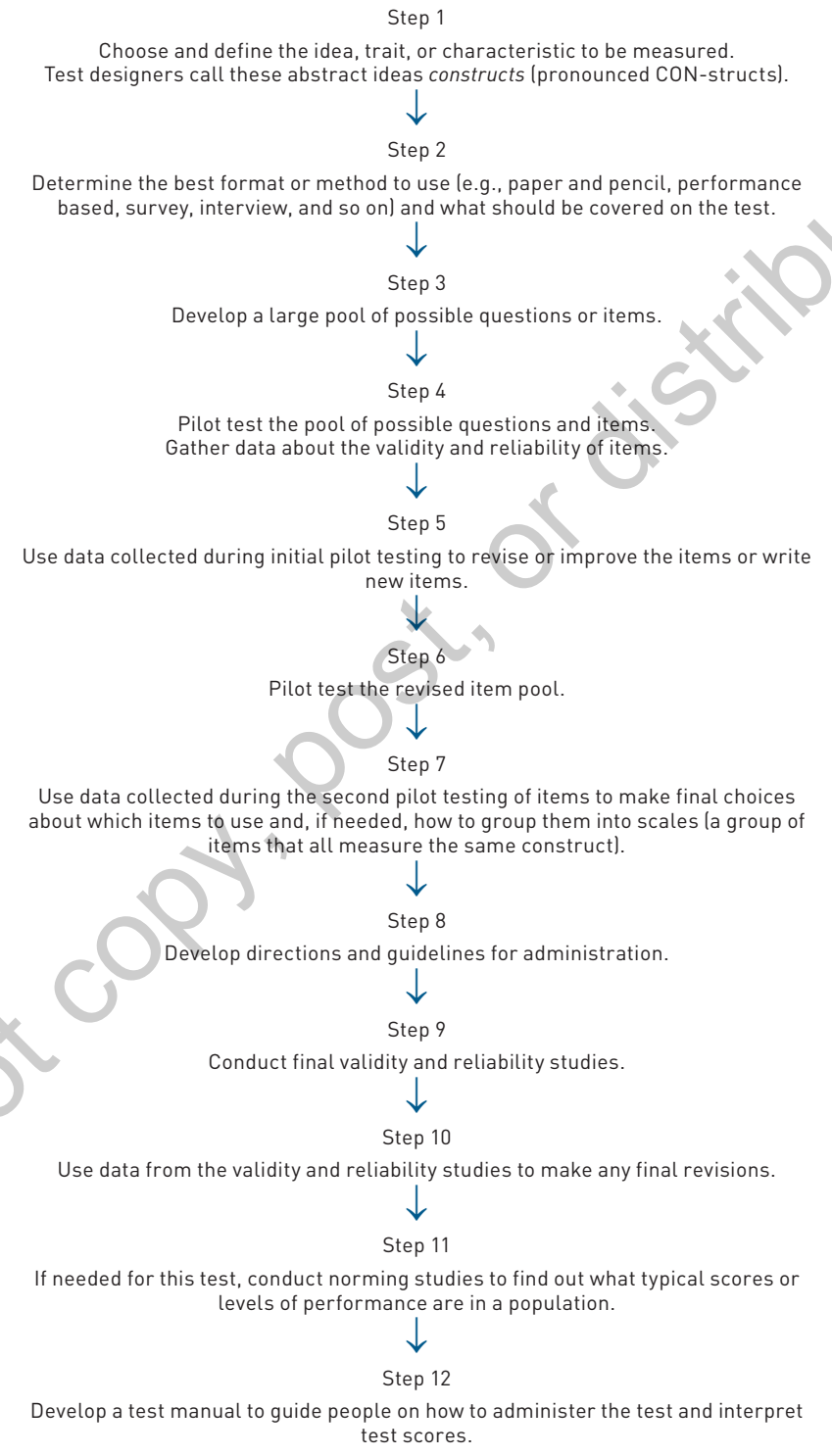
We can suggest several books that are all about the theory and mechanics of test construction, and this is not one of them. So instead, we humbly offer this brief summary of how, in general, a test is designed and the steps in the process. Keep in mind that the process shown here is for standardized tests used for important decision making and, in some cases, for tests used for social science research. A smart teacher using a rubric to grade a book report likely skipped a couple of these steps.

The entirety of the process shown in Table 1.2 is linear; that is, step 2 almost always follows step 1, but within each step there is some evaluation of whether it is time to move on to the next step or repeat a previous step (or even just start all over from scratch). Let's take a look.

So What's New?

Up until the last few decades, the development of almost all tests fell within something called Classical Test Theory (or CTT). The CTT model (and most of this book discusses stuff in the context of that model) primarily looks to increase the accuracy of measuring a test taker's typical score or *true score*, which is a theoretical

TABLE 1.2 ■ A Broad Description of the Steps in the Development of a Standardized Test



average score a person would get if they took the test an infinite number of times. Because there will be some randomness in a person's performance unrelated to the actual level of the trait or construct being measured, the score a person gets on a test (what measurement folks call the *observed score*) is unlikely to be the true score, but might be close to it. The closer that observed score is to the person's true score, the more reliable the test is.

One alternative to Classical Test Theory is Item Response Theory (IRT), which places the emphasis not on the individual's performance and the various sources of random error in the testing situation but instead focuses on the level of reliability in the items themselves. It recognizes that the functioning of an item depends partly on the item and partly on the characteristics of the test taker. For example, on an achievement test, a question will vary in difficulty (and, therefore, in reliability) depending on how much the test taker knows. What's hard for me might not be hard for you.

We'll distinguish between CTT and IRT (as well as some other new approaches) in Chapter 6. All you need to know for now is that, as in almost all disciplines, new ideas and techniques are always being developed, almost always interesting, and surely always ripe for discussion and friendly differences among experts, colleagues, and students as to what's best.

WHAT AM I DOING IN A TESTS AND MEASUREMENT CLASS?

There are probably many reasons why you find yourself using this book. You might be enrolled in an introductory tests and measurement class. You might be reviewing for your comprehensive exams. Or you might even be reading this on summer vacation (horrors!) in preparation and review for a more advanced class.

In any case, you're a tests and measurement student whether you have to take a final exam at the end of a formal course or whether you're just in it of your own accord. But there are plenty of good reasons to be studying this material—some fun, some serious, and some both.

Here's a list of some of the things our students hear when we teach measurement courses. And your instructor might say similar things.

- Tests and Measurement 101 or Introduction to Testing or whatever it's called at your school looks great listed on your transcript. Kidding aside, this may be a required course for you to complete your major. But even if it is not, having these skills is definitely a big plus when it comes time to apply for a job or for further schooling. And with more advanced courses, your résumé will be even more impressive.

- If this is not a required course, taking a basic tests and measurement course sets you apart from those who do not. It shows that you are willing to undertake a course that is (traditionally) above average in regard to difficulty and commitment.
- Basic information about tests and measurement is an intellectual challenge of a kind that you might not be used to. A good deal of thinking is required, as well as some integration of ideas and application. The bottom line is that all this activity adds up to what can be an invigorating intellectual experience, because you learn about a whole new area or discipline. Imagine trying to see something that is invisible. That's what people who use tests are trying to do!
- There's no question that having some background in tests and measurement makes you a better student in the social, behavioral, and health sciences. Once you have mastered this material, you will have a better understanding of what you read in journals and also what your professors and colleagues may be discussing and doing in and out of class. You will be amazed the first time you say to yourself, "Wow, I actually understand what they're talking about." And it will happen over and over again, because you will have the basic tools necessary to understand exactly how scientists reach the conclusions they do.
- If you plan to pursue a graduate degree in education, anthropology, economics, nursing, medicine, sociology, or any one of many social, behavioral, and health sciences fields, this course will give you the foundation you need to move further.
- Finally, you can brag that you completed a course that everyone thinks is the equivalent of building and running a nuclear reactor. (In which Bruce proudly got a C-. He would have done better, but he always pronounced nuclear "new-KYUH-ler.")

TEN WAYS NOT TO HATE THIS BOOK (AND LEARN ABOUT TESTS AND MEASUREMENT AT THE SAME TIME!)

Yep. Just what the world needs—another tests and measurement book. But this one is different (we think, and others have told us so). It is written for you, it is not condescending, it is informative, and it is as simple as possible in its presentation. It assumes you have only the most basic information about testing and the math of measurement—remember mean, median, and mode? Well, that's where the math begins.

There has always been a general aura surrounding the study of tests and measurement that it's a difficult subject to master. And we don't say otherwise, because

parts of it are challenging. On the other hand, millions and millions of students just like you have mastered this topic, and you can, too. Here are a few hints to close this introductory chapter before we move on to our first topic.

1. *You're not dumb.* That's true. If you were, you would not have gotten this far in school. So treat tests and measurement like any other new course. Attend the lectures, study the material, and do the exercises in the book and from class, and you'll do fine. Rocket scientists know how to use this stuff, but you don't have to be a rocket scientist to succeed.
2. *How do you know tests and measurement is hard?* Is this topic difficult? Yes and no. If you listen to friends who have taken the course and didn't work hard and didn't do well, they'll surely volunteer to tell you how hard it was and how much of a disaster it made of their entire semester, if not their lives. And let's not forget—we always tend to hear from complainers. So I suggest that you start this course with the attitude that you'll wait and see how it is and judge the experience for yourself. Better yet, talk to several people who have had the class and get a good general idea of what they think. Just don't base your opinion on one spoilsport's experience.
3. *Form a study group.* This is one of the most basic ways to ensure some success in this course. Early in the semester, arrange to study with friends. If you don't have any who are in the same class as you, then make some new ones or offer to study with someone who looks to be as happy about being there as you are. Studying with others allows you to help them if you know the material better or to benefit from others who know the material better than you do. Set a specific time each week to get together for an hour and go over the exercises at the end of the chapter or ask questions of one another. Take as much time as you need. Find a coffee shop and go there with your study buddy. Studying with others is an invaluable way to help you understand and master the material in this course.
4. *Stay on task and take one thing at a time.* Material about testing and measurement can be tough to understand, especially if you have never heard any of these terms before or thought about any of these ideas. Follow the guidelines mentioned here and talk with your teacher as soon as you find yourself not understanding something or falling behind.
5. *Ask your teacher questions, and then ask a friend.* If you do not understand what you are being taught in class, ask your professor to clarify it. Have no doubt—if you don't understand the material, then you can be sure that others do not as well. More often than not, instructors welcome questions. And especially because you've read the material before class, your questions should be well informed and help everyone in class better understand the material.

6. *Do the exercises at the end of a chapter.* The exercises are based on the material and the examples in the chapter they follow. They are there to help you apply the concepts that were taught in the chapter and build your confidence at the same time. If you can answer these end-of-chapter exercises, then you are well on your way to mastering the content of the chapter.

7. Yes, it's a very old joke:

Q. How do you get to Carnegie Hall?

A. Practice.

Well it's no different with basic statistics. . . .]

Well it's no different with basic statistics. You have to use what you learn and use it frequently to master the different ideas and techniques.

This means doing the exercises in the back of the chapter as well as taking advantage of any other opportunities you have to understand what you have learned.

8. *Look for applications to make it more real.* In your other classes, you probably have occasion to read journal articles, talk about the results of research, and generally discuss the importance of the scientific method in your own area of study. These are all opportunities to look and see how your study of tests and measurement can help you better understand the topics under class discussion as well as the area of beginning statistics. The more you apply these new ideas, the better and more full your understanding will be.

9. *Have fun.* This indeed might seem like a strange thing for you to read, but it all boils down to you mastering this topic rather than letting the course and its demands master you. Set up a study schedule and follow it, ask questions in class, and consider this intellectual exercise to be one of growth. Mastering new material is always exciting and satisfying; it's part of the human spirit. You can experience the same satisfaction here. Just keep your eye on the ball and make the necessary commitment to stay current with the assignments and work hard.

10. *Finally, be easy on yourself.* This is not material that any introductory student masters in a matter of hours or days. It takes some thinking and some hard work, and your expectations should be realistic. Expect to succeed in the course, and you will.



Every now and then, but not often, you'll find steps like the ones you see here. This indicates that there is a set of steps coming up that will direct you through a

particular process. These steps have been tested and approved by the federal agency that oversees staircases, escalators, and small step stools.

The Famous Difficulty Index

For want of a better way to give you some up-front idea about the difficulty of the chapter you are about to read, we have developed a highly secret difficulty index using smileys. (Harvey Ross Ball claimed to have invented smiley faces in 1963, by the way. Please don't tell him we are using them. We cannot handle a lawsuit right now.) This secret code lets you know what to expect as you begin reading. Remember, the MORE smiling faces the EASIER the material!

How Hard Is This Chapter?

Look at Mr. Smiley!

Very hard



Hard



Not too hard, but not easy either



Easy



Very easy



GLOSSARY

Bolded terms in the text are included in the glossary at the back of the book.

Summary

Now you have some idea about what a test is and what it does, what areas of human behavior are tested, and even the names of a few tests you can throw around at tonight's dinner table. But most of all, we introduced you to a few of the major content areas we will be focusing on throughout *Tests & Measurement for People Who (Think They) Hate Tests & Measurement*.

Time to Practice

1. What are some of your memories of being tested? Be sure to include (if you can) the nature of the test itself, the settings under which the test took place, how prepared or unprepared you felt, and your response upon finding out your score.

2. Go to the library (virtually or in person) and identify five journal articles in your area of specialization, such as teaching math or nursing or social work. Now create a chart like this for each set of five.

Journal Name	Title of Article	What Was Tested	What Test Was Used to Test It?
1.			
2.			
3.			
4.			
5.			

- a. Were most of the tests used developed commercially, or were they developed just for this study?
 - b. Which test do you think is the most interesting, and why?
 - c. Which test do you think got the closest to the behavior that the authors wanted to measure?
3. Ask your parent, child, professor, colleague, or classmate what they believe are the most important reasons for testing and what types of tests they can identify.
 4. One of the things we did in this opening chapter was identify five different purposes of tests (see the section, "Why We Test"). Think of at least two other ways that tests might be used, and give a real-world example of each.
 5. Interview someone who uses tests in their work, as either an assessment or a research tool, and try to get an idea of the importance they place on being knowledgeable about testing and what role it plays in their research and everyday professional career. Are they convinced that tests assess behavior fairly? Do they use alternatives to traditional testing? Do they find the results of tests useful for helping students?
 6. Extra credit and extra imagination: Use your favorite search engine and search on five different topics related to testing in general, such as fairness in testing, use of computerized testing, how tests are developed, and so on. Use your imagination and search as broadly as possible. Summarize the results of these searches and propose some directions you think testing might be taking in future activities.