



Teaching For Success®

Quick Answers

No. 809

Q: What Do I Need to Know about Testing?

A: Testing 101 Summary Gives You Quick Answers

by John H. Shrawder

Are you tired of being in the dark about testing and evaluation? Do you agree that grades are extremely important and should represent as accurately as possible a true measure of learning achievement? If so, then we are singing from the same sheet.

Perhaps you've had a similar experience: As a new instructor, I was handed a key to the classroom, and shown the department files that contained tests left by the previous instructor.

Little did I know at the time how poor these tests were. Having no knowledge of how to evaluate tests and test questions, my students suffered. Today, after more study and teaching experience, I know much more and can save you time and mistakes by sharing what I've learned.

If you are serious about teaching and desire to really understand how to better evaluate and grade learning, then spending a few minutes now and again in the quest to understand testing basics is a good investment of your time.

Learn Terms and Take Control

Testing theory and practice has its own special terms that may need some explanation.

If you are new to testing and test terminology, you'll find it helpful to become comfortable with the following terms, especially for use when conversing with fellow faculty, administrators, or for easy reading of reference materials. If you have not yet learned the basics of testing and evaluation for learning, your students are getting short changed.

Below are testing and evaluation terms that are crucial to understanding how to properly and accurately test and evaluate learning:

- Evaluation
- Formative evaluation
- Summative evaluation
- Diagnostic evaluation
- Test
- Measurement
- Performance
- Standard
- Reliability

- Validity
- Norm-referenced
- Criterion-referenced
- Objective-referenced
- Objective question
- Subjective question
- Sample

Here's how to use these terms appropriately. When you feel comfortable with their meaning, you will test students more intelligently, and gain respect as a competent instructor among your fellow instructors, administrators, and students.

Evaluation—Being Systematic

This is the granddaddy, overarching term that means using a systematic process to make sound value judgments.

Formative Evaluation—During

This type of evaluation is used to create a progress report during a course of learning. Quizzes, midterm exams, chapter take-home tests, journal reviews, and portfolio critiques could all be used in a formative evaluation.

Summative Evaluation—The End

One subdivision deserves another. An evaluation that sums up the outcomes of a learning process at the completion of the learning is a summative evaluation. A final exam is usually part of a summative evaluation.

Diagnostic Evaluation—What's Wrong?

This type of evaluation is used to pinpoint learning problems or uncover gaps in knowledge or skills that are preventing the student from progressing.

Test—A Learning Thermometer

Tests are actually instruments that should measure as precisely as possible the outcomes of learning and they are great for cementing learning into long-term memory.

Measurement—By the Numbers

When you measure something, you get the results in numbers and units; for example, a piece of string with a length of 9 inches. Similarly, when you measure learning, you use an instrument that provides a numerical result.

Performance—Getting Results

Performance is about results and outcomes. Performance tests should reveal to learners answers to the questions:

- How well did I learn this?
- How soon can I advance?
- How much better do I know it now than when I started?

These are good performance questions.

Standard—Jump! How High and How Long?

A most important testing concept. When you think about standards in testing you are pondering the question, “How well must my students perform and what will be the measuring stick?” When you create a test without setting or choosing a standard, the results are meaningless. Some good standards are:

- A percentage correct, or a number of items or actions to be completed in a specified amount of time.
- A measure of how accurately work is accomplished compared to a fixed model or list of tasks, perhaps specified in a textbook.
- Comparisons to others taking the same test in the same class or another course section.

Reliability—Repeatable Results Over Time

This term describes how well a test would produce the same results if given to other groups of students at different times and places. A reliable test would yield approximately the same range of scores and share similar mean scores whenever it’s given. When you create your own tests, expect that their reliability will be low.

Validity—Intent is Measured

This term is crucial to understand. Valid tests measure what they are designed to measure. Unless your tests are carefully designed, they will have low validity and fail to measure the knowledge and skills you intend. If your tests lack validity, expect arguments and challenges over correct answers and what the question is really asking. When invalid tests are used to make important grading and competency decisions, tragic errors can result. Valid tests are a must.

Norm-referenced or Grading on the Curve

A norm-referenced test measures the performance standing of an individual in reference to the performance of a group. When you grade on the curve you are norm-referencing the test.

Criterion or Objective-referenced—Mastery Learning

This type of test measures the performance of a student against a defined set of learning tasks or list of learning objectives. Mastery learning uses criterion-referenced tests to evaluate whether a student has mastered an acceptable number of learning objectives.

Objective Question—Easy to Score; Hard to Write

Objective questions are those that can be scored without a detailed analysis of the answer. Multiple-choice, fill-in-the-blank, true-false, matching, and one-word short answer are all examples of objective test questions.




Subjective Question—Easy to Write; Hard to Score

Objective questions are those that are scored by a detailed analysis and repetitive scoring process; whereas, essay test questions are subjective questions. The answers to subjective questions are subject to the examiner's opinion on the correctness of the response.

Sampling—How Many Students; How many “?”

The concept of a representative sample is fundamental to testing. There is simply not enough time to verify that a student can display and apply all knowledge and skills learned. When you test, you sample what the student has learned. The larger the sample, the more faith you and your students can place in the accuracy and fairness of the test.

Need More Details on How to Construct Good Tests?

You're in luck. Teaching For Success has developed an entire Just-in-Time Collection of Quick Studies, monographs, papers, and more on how to teach better and avoid embarrassing mistakes that comprise the TFS Member Online E-library.

Look for the newly revised TFS Quick Study, “Ensure Quality Testing.” It takes you step-by-step through constructing both subjective and objective test questions and tests.


