A TFS Quick Study

How to Test and Evaluate Learning
Section 1. Overview

Student arguments, complaints and disappointments over testing are all a possibility during any course, unless you have developed a sound working knowledge of the evaluation and testing skills required to be a successful instructor. Conflict occurs because students care deeply about tests and test results, since they so profoundly affect their academic careers. When I talk with students, I find poorly written, tricky and ambiguous or irrelevant test questions are at the top of the list of students’ complaints about their instructors. Students deserve good, fair tests, and this QuickStudy can help you create them for your students.

This Teaching For Success QuickStudy™ is designed for you, the busy, caring part- or full-time instructor who desires to develop teaching skills, but must make every minute devoted to professional development count. This self-study course will give you the fundamental knowledge you need to be comfortable and competent in the area of testing and evaluation.

What this course can do for you.
After completing this QuickStudy you should know:

- The purpose of testing.
- Common test and evaluation terms.
- What a test should accomplish.
- How to use a TFS (TPAG) Test Planning and Analysis Grid to check the balance of your tests.
- What question words correspond to various levels of learning achievement as described in Bloom’s taxonomy.
- How to evaluate test questions.
- How to estimate the proper length for a test.
- How to create a Course Testing Plan.
- How to use the TFS Course Testing Planning Grid to create a good course testing plan.
- Resources for more detailed testing information.
Is this Quick Study for me?

The first principle of effective test design is overcoming the notion that testing merely involves writing a set of questions. In truth, it starts with specific, defined course outcomes that are linked to effective lesson plans and based on a sound learning model such as Bloom’s Taxonomy.

Good testing is an outgrowth of a well-organized course that helps students study, understand, apply and retain what has been learned. Therefore, tests should never be constructed as a guessing game for students.

Applicability

This QuickStudy is designed for the instructor in any discipline who is faced for the first time with the tasks of planning, creating or selecting test questions and test types and administering tests in their classrooms.

Time required

You should devote about 60-90 minutes to reading and interacting with this QuickStudy. You will be asked to answer questions, then practice applying this new knowledge to improving a test you created or selected from a test-question bank. You can practice writing test questions in interactive text fields. If time is short, just take it one section at a time.

Instructional design

The learning process in this QC starts with a consideration of the three functions of tests, then moves on to creating a testing plan based on the six learning levels described by Bloom’s Taxonomy.

Rules for test writing are reinforced with sample objective-style questions demonstrating good and bad question construction techniques. Finally, test-length factors are considered, a course testing plan is presented and links to several testing websites are provided.

Main Study Topics

- Testing talk: What do common testing and evaluation terms mean?
- Planning a test: What do you want the test to do?
- Question construction: How do you write or select good, clear, concise and valid test questions?
- Test length: How to estimate the length of a test.
- Course testing plan: How to create a course testing guide.
- Self-test: Check your understanding of the terms, recommendations and principles contained in this QuickStudy.
- References: Book and articles that can provide more testing details.
- Job aides: Blank TFS test design and planning forms.
Section 2. Testing Basics

Since childhood, for many of us, the word “test” or “exam” has been a stress-producing trigger word. Tests are dreaded by some students and faculty alike. Therefore, testing is an emotionally-charged subject for students.

Poor tests and questionable questions will evoke more student disagreements and arguments than any other teaching and learning activity. This QuickStudy will help you improve your tests and your students’ satisfaction with the way you evaluate them.

You may assume that since you have taken a plethora of tests throughout your academic career that you know how to create a good, fair test. If so, you have discovered or are about to discover that testing and evaluation is a complex subject and demands learning a body of knowledge and applying some practical skills.

You have experienced (or if new to teaching, about to experience) the grief that poorly constructed or planned tests can cause students and instructor alike. Without a basic familiarity with good testing principles, you may feel the burden of wondering whether or not your tests adequately evaluate student knowledge and comprehension. This QuickStudy will give you the tools so that you can know for sure that you are doing a professional job in testing and evaluation.

Tests should never be seen by your students as a guessing game or an exercise in interpreting ambiguous questions. Questions should be clear, straightforward and understandable. Good tests begin with good learning objectives.

Learning objectives

One of the most logical ways to look at testing is to recognize that test items are really learning objective statements transformed into questions.

Quality tests depend on quality learning objectives. If you don’t know or can’t define or express precisely what students must know, do or choose, how can you develop accurate tests to measure learning achievement? So job-one is to be able to write or identify learning objectives. If you are new to teaching and have not been given a list of learning objectives or expected outcomes for your course, ask your department head if such a list exists. If you fail to find a prepared list for your course, check the course textbook. Some instructor’s text editions will list learning objectives.

A formal learning objective statement has three parts:

- Starting conditions. (Where do you start?)
- Action. (What will the learner do with the learned knowledge, skills or attitudes.)
- Performance standard. (How well will the learner do it?) Possible time or accuracy standards stated.
How to Get a “A” in Question and Test Selection and Creation

Example three-part Learning Objective

G iven (starting conditions) a PC with a text editor without a spell checker installed, the student will be able to (what learner will do) write a paragraph in a five minutes consisting of a topic sentence, and four supporting sentences.

The paragraph will (performance standard—how well) conform to standard grammar rules with no more than two grammar or punctuation errors and have no more than one word misspelled.

Granted that writing a formal, three-part learning objective (Given, Do, Standard) at this level of detail takes considerable time and effort and therefore, may be an impractical task for many part-time faculty. But the principle of specifying learning outcomes is a sound instructional design and testing principle. To ensure quality testing, it is essential that you have a list of learning objectives or learning outcomes available prior to the preparation of each test.

More than off-the-cuff questions

The first principle of effective testing is dismissing the notion that testing merely involves writing or selecting a bunch of clever or tricky questions. Rather, good testing is an outgrowth of detailed course and lesson planning. Tests should be based directly on what your students must learn and what performance levels must be achieved.

Tests should never be a guessing game for students. Questions should be clear, straight forward, understandable and congruent with the material students must learn. Students should never have to ask, “Will this be on the test?” Your syllabus, lesson plans and learning activities should all make it very clear to students what and how learning will be evaluated.

Not an option

For most of us, we are forced to test by our institutions and federal or state licensing boards or agencies. Paper or computer objective and subjective tests are the most common form of evaluation used today at all educational levels. So, whether to test is not usually an option, but you do have the option to decide how and oftentimes when to test and what the results mean.

Some instructors rely solely on tests extracted from an instructor’s edition of a textbook, but these pre-made tests are geared toward the average student taking an average version of the course. Therefore, they may not ensure quality testing unless the questions are evaluated and selected according to a criteria based on a specific course and student group.
Since you are an individual and unique teacher, chances are your classes reflect your specific strengths and experiences. Your tests should therefore reflect your special approach, and often, the only way to accomplish this is to create tests from your own test questions.

Whether you are faced with the task of writing test questions from scratch, selecting questions for a test from a bank of publisher-supplied test questions or choosing a test from those offered by your textbook publisher, the information in this *QuickStudy* will help you evaluate your students with a high quality test.

**Your turn**

Describe in a few brief sentences in the text field below the most pressing problems that you face with testing and evaluation.

**Since you are an individual and unique teacher, chances are your classes reflect your specific strengths and testing style. The best way to express your unique evaluation style is to learn how to create good, fair tests.**

Why is a list of learning objectives or learning outcomes essential to have before a test is created? Write your answer in the field below.

Name the three parts of a formal learning objective statement?

**Part 1.**

**Part 2.**

**Part 3.**

Answers?

Click here to go to answers.
How’s Your Testing Talk?

Testing theory and practice has its own special terms that may need some getting used to. If you are new to testing and test terminology, you’ll find it helpful to be comfortable with the terms when conversing with fellow faculty, administrators or for easy reading of reference materials. These are testing and evaluation terms you should know: (term links are active):

- 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 are the simplified, working definitions that we use at Teaching For Success: Use these terms appropriately and you will gain respect as a competent instructor!

Evaluation—This is the granddaddy term that means using a systematic process to make a sound judgments about the value of something.

Formative evaluation—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—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—And, there is even another subdivision. This term describes a type of evaluation that is all too often neglected in the higher education classroom. It is used to pinpoint learning problems or uncover gaps in knowledge or skills that are preventing the student from progressing.
A TFS Quick Study  

How to Get a “A” in Question and Test Selection and Creation

**Test**—In testing circles, tests are referred to as **instruments** that measure how well a learner performs.

**Measurement**—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**—Now, this is a term that really piques the interest of an Educational Behaviorist. Performance is about results: How well? How soon? How much better? These are good performance questions. To be meaningful, performance should be clearly defined and measured against a standard.

**Standard**—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? You can create a test, but without setting or choosing a standard, the results are meaningless. Standards, for example, may be expressed in percentage correct, or as a number or items or actions to be completed in a specified amount of time or how accurately work is accomplished compared to a fixed model or list of tasks, or compared to others performing the same action in the same class or in all course sections.

**Reliability**—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 you can expect that their reliability will be low. The same test may yield varying scores from use to use.

**Validity**—This term is crucial to understand. Valid tests measure what they are designed to measure.
How to Get a “A” in Question and Test Selection and Creation

The importance of validity

Unless your tests are carefully planned, designed and constructed, they will have low validity and fail to measure the knowledge, skills or attitudes that you intend. If your tests lack validity, expect arguments and challenges over correct responses 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 good, fair and accurate tests of the knowledge and skills to be assessed.

How to Maximize Validity

- Ask a peer, friend or spouse to review the questions and identify ones that are confusing or ambiguous.
- Check question construction according to the construction principles found in the “Question Construction” section of this course.
- Use a test construction grid to match questions with learning objectives. Eliminate any superfluous questions.
- Ensure students know precisely how to answer each question as well as the scoring and grading criteria.

Norm-referenced—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—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—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—Objective questions are those that must be scored by 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.
Sample—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 better the sample, the more faith you and your students can place in the accuracy and fairness of the test.

Your turn

Check Your Test Talk

1.) Which testing-talk term describes a systematic process of making value judgements?

2.) A final exam is an example of which type of evaluation?
   1. Formative.
   2. Summative.
   3. Diagnostic.

3.) A test that measures exactly what it was designed to measure has what characteristic?
   1. Subjectivity.
   2. Objectivity.
   3. Reliability.
   4. Validity.

4.) A question that features responses that can be rapidly and accurately graded by anyone is an example of a:
   1. Subjective question.
   2. Objective question.
   3. Performance question.
   4. Valid question.

5.) A teacher grading a test finds that the average score is 78 percent. The instructor then looks for groups of scores higher and lower than the average and assigns these groups “A”, “B”, “D” and “F”, with a “C” being from 80-76 percent. This is an example of an instructor using which kind of referenced test?
   1. Norm-referenced.
   2. Objective-referenced.
   3. Criterion-referenced.

6.) A test that yields a similar range of scores class after class would have what level of reliability?
   1. High.
   2. Medium.
   3. Low.

Click to Check Your Answers
A Quick Study
How to Get a “A” in Question and Test Selection and Creation

Section 4. Test Planning

Abraham Lincoln was quoted as saying that if he had only four hours to chop down a tree he would spend at least three hours sharpening his axe. He knew a dull tool wastes both time and effort.

When you spend time in test planning, you correctly treat the test as a critical tool that you will hone to perfection before attempting to accurately measure student learning.

Planning steps
The test planning process consists of the following steps:
- Define the overall purpose of the test.
- Identify the applicable learning objectives or outcomes that define the content of the test.
- Classify each learning objective or learning outcome to one of Bloom’s learning levels.
- As you select or write questions to check each learning objective, complete a question analysis grid analyzing the question distribution using Bloom’s six learning levels.

Purpose
Decide if the test that you are planning is a pretest, a norm-referenced chapter test (how are individual students learning compared with the class as a whole), a diagnostic test to determine why students did not reach a specific learning objective, or a summative test such as a criterion-referenced final test to affirm that each student has mastered (pass-fail) the necessary learning objectives.

Organizing learning is at least half of the challenge of effective test planning. Benjamin Bloom’s Cognitive Domain Taxonomy of Educational Objectives describes six major levels of cognitive learning. It’s a terrific way to organize the learning tasks in your course.

Unless you already use another learning model that works for you, we recommend that you give this organizational scheme a try. It can greatly simplify test planning.

A handy chart of Bloom’s Taxonomy of Learning has been created for you that includes the names of each of the six levels of learning coupled to a description of what a student should be able to do when competent at that level.

What’s Bloom’s Taxonomy all about?
More information about Bloom’s Taxonomy is available at (http://www.coun.uvic.ca/learn/program/hndouts/bloom.html) Click on the link URL to visit this site.
**Bloom’s Learning Levels**

<table>
<thead>
<tr>
<th>Level</th>
<th>What a student should be able to do:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge</td>
<td>Recall facts such as names, dates, events, places, etc.</td>
</tr>
<tr>
<td>Comprehension</td>
<td>Understand facts, grasp meaning, put knowledge into a new context.</td>
</tr>
<tr>
<td>Application</td>
<td>Apply knowledge and skills to solve problems, to use or demonstrate applications of rules, procedures and processes.</td>
</tr>
<tr>
<td>Analysis</td>
<td>Reduce to smaller increments, to see patterns, to organize parts, etc.</td>
</tr>
<tr>
<td>Synthesis</td>
<td>Fit increments together to produce new material, to generalize from a body of facts.</td>
</tr>
<tr>
<td>Evaluation</td>
<td>Constructively critique the work of others, to verify and judge the value of information or opinions, etc.</td>
</tr>
</tbody>
</table>

Why will using this chart help you plan a better test? Without careful planning and discipline many instructors tend to choose test questions that focus almost exclusively on the lowest levels of learning. Without realizing it, many inexperienced test creators will ask a disproportionate number of questions in the Knowledge, Comprehension and Application levels. If you desire to teach and test for higher-order thinking skills, then it is important to include questions at the Analysis, Synthesis and Evaluation levels.

In addition, writing a Knowledge-level question is generally much easier than creating a Synthesis-level question. Therefore, over-committed and time-pressed instructors will gravitate toward producing the highest number of test questions per hour as possible.

You will use the information in this chart when completing the next step in test planning, which is to complete the TFS Test Planning and Analysis Grid, (TPAG)

To help you organize a great deal of information in a small space, it’s helpful to create information codes such as the set of question type codes we have created to work with the TFS Test Planning Grid that you will see on page 13. The table below lists the seven most commonly used question types and an abbreviation for each question type to use in the test planning process.

<table>
<thead>
<tr>
<th>TFS Q Type Codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>MC</td>
</tr>
<tr>
<td>TF</td>
</tr>
<tr>
<td>MAT</td>
</tr>
<tr>
<td>ES</td>
</tr>
<tr>
<td>SA</td>
</tr>
<tr>
<td>COM</td>
</tr>
<tr>
<td>IE</td>
</tr>
</tbody>
</table>
The TFS Test Planning and Analysis Grid (TPAG)

The above Test Planning and Analysis Grid developed at Teaching For Success, maps four crucial components for planning a valid test:

- Performance levels.
- Learning objectives.
- Number of questions.
- Question type.

Using this or a similar table to plan your tests will help you create a much more balanced, focused and valid test. Begin by estimating the percentage of the test questions you will devote to each of Bloom’s cognitive performance levels. In the example above, the target percentages have been entered below each performance level in the light gray shaded cells.

Down the left-hand side of the Grid, list the learning objectives to be tested. Then using the Q Type Codes fill in the cells with question type, question quantity and the Q focus or Question Focus.

In the sample planning grid above, note: there are three learning objectives entered from Chapter 1 or the sample course textbook. These objectives are associated with seven questions separated into Bloom’s performance or knowledge levels.
How to Get a “A” in Question and Test Selection and Creation

The TFS Test Planning and Analysis Grid (TPAG)

Course: Introduction to Amateur Radio, Electronics 101
Type, Basis, Purpose and Scope: Unit Test, Norm-referenced, Assess Class Progress, Chapters 1-3

Performance Levels

<table>
<thead>
<tr>
<th>Learning Objectives</th>
<th>Knowledge Target (30%)</th>
<th>Comprehension Target (20%)</th>
<th>Application Target (30%)</th>
<th>Analysis Target (5%)</th>
<th>Synthesis Target (5%)</th>
<th>Evaluation Target (10%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ch 2. “Fundamental Radio Theory—DC Circuits”. Apply Ohm’s law and calculate values of voltage, current and resistance. Total Q: 3 = 15%</td>
<td>#Q; Type 3-MC</td>
<td>Q Focus 1-3.</td>
<td>Quantity 3</td>
<td>#Q; Type</td>
<td>Q Focus 1. Transistor; 2. Diode; 3. SCR; 4. LED; 5. IC.</td>
<td></td>
</tr>
<tr>
<td>Ch 3. “Solid State Devices”. Name 10 common solid state devices; select a device to perform a specified task. Total Q: 10 = 50%</td>
<td>#Q; Type 5-MC</td>
<td>Q Focus 1. Transistor; 2. Diode; 3. SCR; 4. LED; 5. IC.</td>
<td>Quantity 3</td>
<td>#Q; Type 5-MC</td>
<td>Q Focus 1. Transistor; 2. Diode; 3. SCR; 4. LED; 5. IC.</td>
<td></td>
</tr>
<tr>
<td>Test Total Questions = 20</td>
<td>7 MC = 7</td>
<td>1 SA + 2 TF = 3</td>
<td>1 Mat + 8 MC = 9</td>
<td>1 Essay = 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage in each Level</td>
<td>7/20 = 35%</td>
<td>3/20 = 15%</td>
<td>9/20 = 45%</td>
<td>0/20 = 0%</td>
<td>0/20 = 0%</td>
<td>1/20 = 5%</td>
</tr>
</tbody>
</table>

At this point in this example planning process, the instructor assigns questions for the learning objectives found in chapters two and three of the course textbook.

For Chapter Two, the instructor has decided to concentrate the questions asked exclusively at the Application level. She plans to ask three multiple-choice questions concerning the application of Ohm’s Law to calculating the voltage and current in DC series and parallel circuits. In Chapter Three she decides to ask five questions at the Knowledge level and five at the Application level.

Again, multiple-choice is seen as the desired question type for these questions. The bottom two rows of the table are summary rows that help the instructor now see the overall balance in this 20-question unit test.
The TFS Test Planning and Analysis Grid (TPAG)

**Course**: Introduction to Amateur Radio, Electronics 101

**Type, Basis, Purpose and Scope**: Unit Test, Norm-referenced, Assess Class Progress, Chapters 1-3

<table>
<thead>
<tr>
<th>Performance Levels</th>
<th>Knowledge Target (30%)</th>
<th>Comprehension Target (20%)</th>
<th>Application Target (30%)</th>
<th>Analysis Target (5%)</th>
<th>Synthesis Target (5%)</th>
<th>Evaluation Target (10%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning Objectives</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ch. 1. “Introduction”</td>
<td>#Q: Type 2-MC (This means two multiple-choice question.)</td>
<td>Q Focus 1. When was amateur radio first officially sanctioned by U.S. law.</td>
<td>#Q: Type 1-SA; 2-TF</td>
<td>Q Focus 1. Purpose of this radio service.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State the purpose and year of the first laws that recognized the value of Amateur Radio service. List applications of AR. Make an evaluation as to the value of AR to civil defense.</td>
<td>Quantity 2</td>
<td>Quantity 3</td>
<td>Quantity 1</td>
<td>Quantity 0</td>
<td>Quantity 0</td>
<td>Quantity 1</td>
</tr>
<tr>
<td>Total Q: 7 = 35%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ch 2. “Fundamental Radio Theory—DC Circuits”. Apply Ohm’s law and calculate values of voltage, current and resistance.</td>
<td>#Q: Type 3-MC</td>
<td>Q Focus Q 1-3. Apply Ohm’s law to DC circuits.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quantity 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Q: 3 = 15%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ch 3. “Solid State Devices”. Name 10 common solid state devices; select a device to perform a specified task.</td>
<td>#Q: Type 5-MC</td>
<td>Q Focus 1. Transistor; 2. Diode; 3. SCR; 4. LED; 5. IC.</td>
<td>#Q: Type 5-MC</td>
<td>Q Focus 1. Transistor; 2. Diode; 3. SCR; 4. LED; 5. IC.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quantity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Q: 10 = 50%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Q Type Codes and Totals**

- 15 MC Multiple-choice
- 2 TF True-false
- 1 MAT Matching
- 1 ES Essay
- 1 SA Short-answer
- 0 COM Completion
- 0 IE Interpretive exercise

**20 Total Questions**

<table>
<thead>
<tr>
<th>Total Questions = 20</th>
<th>7 MC = 7</th>
<th>1 SA + 2 TF = 3</th>
<th>1 Mat + 8 MC = 9</th>
<th>1 Essay = 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage of Total = 100</td>
<td>7/20 = 35%</td>
<td>3/20 = 15%</td>
<td>9/20 = 45%</td>
<td>0/20 = 0%</td>
</tr>
</tbody>
</table>

The planning grid (shown above) has been completed for all three chapters; now analyze the red numbers.
What’s your final analysis?

Looking at the final results of the TPAG by chapter, you may notice the heavy emphasis on Chapter Three material at the expense of Chapter Two learning. Analyzing by Performance Level, one discovers that 45 percent of the questions are concentrated at the Application level perhaps owing to the lack of questions at the Analysis and Synthesis levels.

With only one question focused on the upper three Performance levels, this may indicate a need to create some additional questions focused on using these higher-order thinking skills. The final decision on this is up to you. The TPAG helps to paint a clear picture of the scope and depth of the planned test.

Of course, no test will be perfectly balanced in all aspects, but by taking the time and making an effort to create and review a TPAG or a table of your own design, you will be able to spot serious imbalances in the:

- Number of questions targeting each of the six learning levels.
- Number or questions aimed at each chapter.
- Frequency of each question type used.

The payoff for using such a chart is a higher validity test that you can stand behind with confidence.

Planning Test Length

Now that you have planned and analyzed the test content for balance, check your test for length. To become highly competent at estimating the best length test to give to a specific student group taking a specific course requires trial-and-error learning and practical experience.

However, a place to start is with the following rule: allow between 30-seconds to a minute for each objective-type question depending on the complexity and knowledge level the multiple-choice question is examining. When considering the entire spectrum of possible objective questions, consider that true-false questions take the least time and the interpretive exercise requires the most time to complete.

Also matching questions often require more time to respond as five or six matches must be made for each question. And the same consideration for extra time applies to questions that involve calculations data interpretation. The more involved the calculations or data analysis, the more time should be allotted—perhaps twice the time you would need.

Properly constructed questions will help students progress more quickly through the test allowing you to increase the sample size, and therefore, you help students maximize their performance.

If you add five minutes for administrative duties such as handing out the test and announcing the directions for the test, you should allow approximately 30 minutes.
How to Get a “A” in Question and Test Selection and Creation

A bit shorter is better

If you plan a test that is about 10 minutes shorter than the testing period, you can allow extra time at the end of the test if you observe that a significant number of students have not yet finished. The use of a simple “Time Analysis Grid” will help you plan tests that challenge but not panic students.

<table>
<thead>
<tr>
<th>Q Type</th>
<th>Q Number</th>
<th>Est. Time Each</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>MC</td>
<td>15</td>
<td>60 sec.</td>
<td>900 sec.</td>
</tr>
<tr>
<td>TF</td>
<td>2</td>
<td>30 sec.</td>
<td>60 sec.</td>
</tr>
<tr>
<td>MAT</td>
<td>1</td>
<td>60 sec.</td>
<td>60 sec.</td>
</tr>
<tr>
<td>SA</td>
<td>1</td>
<td>60 sec.</td>
<td>60 sec.</td>
</tr>
<tr>
<td>ES</td>
<td>1</td>
<td>600 sec.</td>
<td>600 sec.</td>
</tr>
<tr>
<td></td>
<td>Totals</td>
<td></td>
<td>1680 sec.</td>
</tr>
<tr>
<td></td>
<td>Divide by 60</td>
<td></td>
<td>28 min.</td>
</tr>
</tbody>
</table>

How many questions should I ask?

Students can usually complete one multiple-choice question or finish two true-false items every minute. The actual time your students need to complete a test depends on many factors so use these estimates until you obtain some time data from your tests.

It is common for the fastest students to complete a test in half the time required by the slowest students.

It is wise to be conservative in your time estimates, especially with a new test. Unless you specifically need to administer a test that pressures students to perform to a time standard, adjust the length to allow students to work at a normal pace. This will ensure a more representative measurement of their level of understanding.

Tests containing lengthy computational problems or essay questions should include time estimates in the test directions to assist students in better planning their test-taking strategy.

By carefully deciding when to test, how often to test, what question types to use, the number of questions to include, and the purpose of the test, you will feel more confident in your ability to accurately test learning. And, your students will feel more confident in a grade that accurately reflects their achievement—a win-win situation for all.
How to Get a “A” in Question and Test Selection and Creation

Your turn

Check Your Test Planning Knowledge

1.) Step one of test planning includes which of the following activities?
   1. Define the purpose of the test.
   2. Assign a Performance level to each question.
   3. Track the number of questions created.
   4. Analyze the question type percentages.

2.) The second performance level of Bloom’s Taxonomy of Cognitive Learning is described by which of the following terms?
   1. Formative.
   2. Analysis.
   3. Comprehension.
   4. Application.

3.) What information is entered in the cells forming the left-most column of the TPAG?
   1. Question type.
   2. Chapter and associated learning objectives.
   3. Analysis figures of the question usage percentage.
   4. Number of multiple-choice questions used.

4.) What was the main content focus imbalance revealed by the TPAG in the test planning sample?
   1. Chapter one may be over emphasized.
   2. Fifty percent of the questions cover Chapter three.
   3. Too many true-false questions are used.
   4. The total number of test questions should be increased.

5.) What was the Performance Level imbalance revealed by the TPAG in the test planning sample?
   1. Too many Knowledge level questions asked.
   2. Too many Application level questions asked.
   3. Too few Application level questions asked.
   4. Too many Synthesis questions asked.

6.) What is the recommended way to estimate the approximate length of a test that you are creating?
   1. Fill out a TPAG.
   2. Give the test to a colleague to take.
   3. Use two-minutes per question as a good rule-of-thumb.
   4. Fill out a Test Time Analysis Grid.

7.) You notice that the top students in the class finish a test in one-half the time required by the slowest students. What is indicated about the test length?
   1. It’s too long.
   2. It’s about right.
   3. It’s too short.

Click to Check Your Answers
Section 5. Question Construction

Testing and measurement researchers are quick to assert that each question type can test understanding at all learning levels if properly constructed. Even though creating a true-false question that tests at the evaluation level may be possible, it is better to go with the question type most suited to the task (see chart to follow).

With at least eight specific categories of test questions to choose from, deciding which type to use can be quite complex. All too often the choice of question type is based on expediency of either construction or scoring or both rather than a consideration of the advantages and disadvantages of each type. Learning how to construct valid and fair test questions is an important skill for higher-education instructors and teaching assistants. A good test results from the selection of well-written test questions.

The list below includes the most popular type of test questions:
- MC  Multiple-choice and true-false.
- MAT  Matching.
- FB  Fill-in-the-blank (completion).
- SA  Short answer (restricted response).
- ES  Essay (open response).
- IE  Interpretive exercise.

The “Question Type Advantages and Disadvantages” chart will help you select the best question type for your purpose.

Learning how to construct valid and fair test questions is an important skill for higher-education instructors and teaching assistants. A good, fair test results from the selection of well-written test questions.

<table>
<thead>
<tr>
<th>Q Type</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>MC</td>
<td>MC and True-false are highly flexible and especially work well to test the lower learning levels such as knowledge, comprehension and application. Four-choice questions are the best compromise. Can be machine scored.</td>
<td>While MC questions can be written to test Synthesis, Analysis and Evaluation levels, these are difficult to write. An essay or Short-answer question may be a better choice for testing at the higher learning levels.</td>
</tr>
<tr>
<td>MAT</td>
<td>Matching allows you to quickly test a range of knowledge with fewer questions—great, for example, in testing terminology.</td>
<td>More time consuming to create than multiple-choice and not well suited for Analysis, Synthesis, and Evaluation levels.</td>
</tr>
<tr>
<td>FB and SA</td>
<td>Fill-in or Short-answer questions are very easy to construct. Great for quizzes and are better indications of knowledge retention because the correct answer is not visible.</td>
<td>Usually not machine compatible. Compatible with all learning levels. Easy to construct, time consuming to grade with large classes.</td>
</tr>
<tr>
<td>ES</td>
<td>Student selects facts, organizes information, draws conclusions, uses higher learning skills such as analysis and synthesis. Can be used in open-book and take-home and group formats.</td>
<td>Scoring criteria must be created and added to question. Time consuming to grade accurately and fairly—favors the students with better writing skills.</td>
</tr>
<tr>
<td>IE</td>
<td>Compatible with Application, Synthesis, Analysis and Evaluation levels. Tests complex areas.</td>
<td>Requires more time effort to construct good questions than essay, but it’s easier to score.</td>
</tr>
</tbody>
</table>
How to Get a “A” in Question and Test Selection and Creation

Learning-level question construction chart

Use the terms in the right-hand column in the chart below to help you construct a question that will test learning at the level listed in the left-hand column.

<table>
<thead>
<tr>
<th>Learning Level</th>
<th>Initial Term in Question Statement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge</td>
<td>Describe, identify, label, select, name.</td>
</tr>
<tr>
<td>Comprehension</td>
<td>Predict, estimate, distinguish, summarize.</td>
</tr>
<tr>
<td>Application</td>
<td>Calculate, solve, show, use, change.</td>
</tr>
<tr>
<td>Analysis</td>
<td>Illustrate, diagram, divide, outline.</td>
</tr>
<tr>
<td>Synthesis</td>
<td>Categorize, organize, rearrange, combine.</td>
</tr>
<tr>
<td>Evaluation</td>
<td>Justify, interpret, compare, contrast.</td>
</tr>
</tbody>
</table>

Your turn: Construct a multiple-choice question

In the field below, write a practice, four-choice, multiple-choice question for testing at the knowledge learning level. This will make the discussion on construction do’s and don’ts more meaningful. Remember at the Knowledge Level a student might be asked to: recall facts such as names, dates, events, places, etc. See the “Learning Level Question Construction Chart” above for a list of key question words to help get you started writing.

Do your questions pass the test?

<table>
<thead>
<tr>
<th>Yes/No</th>
<th>Only one problem or issue is stated per question unless it is an complex essay question.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes/No</td>
<td>The stem is a complete sentence in the form of a statement or a question.</td>
</tr>
<tr>
<td>Yes/No</td>
<td>The question or distracters don’t provide clues to the correct answer.</td>
</tr>
<tr>
<td>Yes/No</td>
<td>The stem is expressed in a positive manner not a negative. (Example of a negative construction: Which of these is not the right answer?)</td>
</tr>
<tr>
<td>Yes/No</td>
<td>The correct response answer is clearly related to the stem with no ambiguity.</td>
</tr>
<tr>
<td>Yes/No</td>
<td>The stem and distracters are free of grammar and spelling errors.</td>
</tr>
<tr>
<td>Yes/No</td>
<td>In multiple-choice items, all choices are about the same length; the longest-answer-is-correct problem has been avoided.</td>
</tr>
<tr>
<td>Yes/No</td>
<td>My students are familiar with this type of question and have practiced answering them in a review exercise.</td>
</tr>
</tbody>
</table>

What about humor?

Some instructors recommend adding humorous distracters to an occasional multiple-choice question or a humorous question somewhere in a test to alleviate test anxiety. Others
feel adding humor is a waste of the students’ time, interrupts concentration and trivializes the testing process.

Compare the multiple-choice question that you wrote on page 20 to the following examples of multiple-choice questions. As you go through the examples, you will sharpen your ability to detect poor-quality questions and add to your knowledge of how to improve them.

**Multiple-choice question considerations**

A multiple-choice question begins with the *stem*, a question or a statement. The stem is followed by one correct response and typically one-to-four distracters or incorrect answers.

**Number of choices**

The more choices you use in a question the less the chance of guessing the right answer, but it’s more difficult to create plausible choices that don’t just give away the correct response.

All of the above; none of the above

Test experts don’t seem to agree on the use or avoidance of “all” or “none of the above” in a multiple-choice question.

The TFS recommendation is to use them sparingly if at all. The “all” or “none” terms invite arguments when the test taker invariable thinks of an exception. It’s very poor practice to arbitrarily add “all” or “none of the above” to every test question. Research shows that in most situations, other than when numerical or symbolic answers are required, a four-answer, multiple-choice question is the best compromise.

**Question diagnostics—become a question doctor**

Several questions will follow. Evaluate the questions as you read them for construction or interpretation problems. See if you can spot a problem or weakness in the question and determine an appropriate fix.

Q). Which Right Path in Buddhism can be compared to an athlete being in the zone? (stem)

a. Right effort. (distracter)

b. Right mindfulness. (correct response)

c. Right thought. (distracter)

d. Right action. (distracter)

Did you spot the problem?

The stem or statement of this question is just ambiguous enough to be argued by many students. The ambiguity lies within the question, regarding the definition of “zone.”

In addition, if a student has never heard of the phrase “in the zone” or formed a working definition of this concept, he or she would have no basis to choose a response and would be left having to guess.
How to Get a “A” in Question and Test Selection and Creation

A fix would be to ask the student to define the “Right Path” in a question or to write a short essay detailing how they may have experienced “Right Path” in their life or ensure “zone” has been carefully defined in the course learning activities.

**Question or puzzle?**

**Ex 1.) Which function of a speech introduction enhances the presenter’s credibility?**

a. Gain and maintain attention.  
b. Relate topic to speaker.  
c. Relate topic to audience.  
d. Preview the body of the speech.

Perhaps you found yourself reading this question and its choices several times, and wondering what is being asked by this question. If you reacted with, “I don’t get it? Is this a trick question surreptitiously asking the reader to differentiate between a speech and a presentation?” we would be thinking alike. There is a place for puzzles, but generally not on a graded test.

Whatever the case, this type of question should be purged from your tests. Good multiple-choice questions start with the writing of clear, concise stems that clearly communicates to a wide-range of readers the exact focus of the question to be answered or task to be completed.

**Stem clarity is poor and mixed distracters used**

**Ex 2.) Charles Darwin is credited for developing which of the following?**

a. Theory of devolution.  
b. Theory of evolution.  
c. Theory of revolution.  
d. Law of evolution.

**Here, crystal clear stem and choices**

**Ex 3.) Charles Darwin is credited for developing which of the following theories?**

a. Radiation.  
b. Evolution.  
c. Mitosis.  
d. Relativity.

**Discussion**

Which version is better? The stem of example Ex 1. is not a complete sentence. What exactly does, “which of the following” mean? This wording opens up the possible correct responses to be almost anything related to Darwin. One has to read and understand the choices to know what the question is asking. Therefore, this is a poorly constructed question.

Going beyond the stem construction problem, which question one should use depends on the learning objective that is related to this question (Remember each question...
How to Get a “A” in Question and Test Selection and Creation

asked should correlate to a learning objective or learning outcome).

Ex 1.) Charles Darwin is credited for developing which of the following?
   a. Theory of devolution.
   b. Theory of evolution.
   c. Theory of revolution.
   d. Law of evolution.

If the learning objective of Ex 1. asks students to discriminate between a “law” and a “theory,” then the a.-d. choices shown above are valid. It’s the stem that needs rewriting. It should ask the student to differentiate between a scientific law and a scientific theory.

How might this be accomplished? I would use a matching format. For example, “In science there are both laws and theories. Identify each of the following terms as either a law or a theory. Place an “L” if it is a law and “T” if it is a theory.

___ Evolution.  ___ F=MA.
___ Gravity.  ___ Ohm’s.
___ Big Bang.  ___ String.
___ Speed of light. ___ Thermodynamics.

However, if the learning objective of Ex 1. is about Darwin and recognition of the theory of evolution that he developed, the question could be stated more clearly by writing EX 1a:

EX 1a.) Charles Darwin is credited for developing which of the following theories?
   a. String.
wrong in the learning process. This can be useful in modifying the teaching and learning activities to avoid the error or misconception.

**True-false**

True/false questions can also be called “either/or” type questions. Either these questions are so easy that everyone gets them right or they are so equivocal that the better students miss them because they know too much.

Good true/false questions are very difficult to write because the statements must be either completely right or wrong—no exceptions, and it seems that the sharper students are always willing to challenge the answer.

**High probability of guessing the correct response**

This type of question is also the favorite of students who prefer to guess answers since the probability of getting the question right by chance alone is 50 percent.

Therefore, scores on true/false tests tend to run considerably higher than tests constructed of other questions; a scoring range of 80 to 100 percent is typical. Finally true/false tests are useless as diagnostic instruments, because they yield no clues as to why the student missed the question.

If a 50 percent guess rate is acceptable, then a simple true-false question could be constructed such as:

**Q). T. F. Charles Darwin is most noted for developing the theory of evolution.**

This is an example of a good, clear, straight-forward true-false question with “T” being the correct response.

**Matching**

Matching tests consist of paring factual information in one column with factual information in another. There may be either **perfectly** or **imperfectly** matched sets of information.

The **imperfect sets** are more difficult since one list provides more choices than the other and the last match cannot be made by a process of elimination. Some of the types of information that form good matching sets are: events and dates, persons and achievements, definitions and terms, principles and applications, and rules and examples.

The columnar format of the matching test allows a large amount of material to be measured in a short time. The test elements are somewhat difficult to design well because you must offer the student several possible matches that are wrong but look plausible.

For this reason, matching questions must center on specific areas of knowledge, and you must take care not to include knowledge that was omitted from the course or taught very superficially.
How to Get a “A” in Question and Test Selection and Creation

Placing the information in each column in a logical order will assist students in locating all possible matches. Be sure to include instructions that explain the matching criteria. Do you see any problems with the matching question below?

Q). Match the animal with its characteristic by placing the number of the animal in the blank provided.

1. Horse ___A. Has a mane.
2. Lion ___B. Has claws used for hunting.
4. Eagle ___D. Eats plant material.
5. Duck ___E. Lives in family units.

Analysis

This matching question has poor validity because the writer has not explored all the possible matches that a test-taker might make.

Horses and lions have manes; bears and eagles eat a diet high in fish; and bears, ducks, and even horses live in family units at least part of the time as they are growing up.

The question could be repaired by stating the directions that there may be more than one match for each characteristic on the right.

Question revision

Match the animal with its characteristics by placing the number of the animal in the blanks provided.

1. Horse ___, ___A. Has a mane.
2. Lion ___, ___B. Has claws used for hunting.
3. Bear ___, ___C. Eats a diet high in fish.
4. Eagle ___, ___D. Eats plant material.
5. Duck ___, ___E. Lives in family units.

Fill-in-the-blank

Leave one blank for each concept being tested per question. In other words, use only one blank per question. And resist the temptation to adjust the blank length to suggest the correct response.

This type of question is excellent for a pop quiz at the start of class. Several questions can be asked and the answers graded by exchanging papers all within five minutes of class time.

Here is an example of a poorly constructed fill-in-the-blank question.

Q). Fill in the blanks with the best term to complete the sentence.

The only certain measure of success is to render more and better service than is expected of you.

—O.G. Mandino
A mouse that _____ to perform a task by being ________ after doing the task is exhibiting ________ conditioning.

**Question revision**

A properly constructed fill-in question only asks for one term to be identified.

Q). Fill in the blank with the best term to complete the sentence.

A mouse that learns to perform a task by being rewarded after doing the task is exhibiting ________ conditioning.

**The test writer’s mind set**

Success Tip: As you write or select questions, adopt the test-taker’s mentality. The prime question in each test-taker’s mind is, “What am I being asked? Is this a straight-forward questions or one with a trick?” And avoid convoluted “not” questions. This construction tricks many students.

Take extra care to create or select clear, honest questions. This is not an effort that will dumb-down the test, but it will ensure each question has good validity, and it will ensure you actually test what you intend to test.

**Essay**

Choose essay questions to measure complex learning such as: the organization and expression of ideas, the integration of concepts, the solution of problems, and the student’s ability to evaluate concepts, processes, or products.

**Limitation**

Essay questions are time consuming to answer, and therefore, an essay test can only cover a limited range of knowledge.

**Benefits**

But essay questions can be used to probe a great depth and breadth of understanding. The effect of guessing on the test outcome is minimized. Also, these questions give more control to the student by allowing a variety of correct responses to a single question. The essay format lends itself well to take-home and group-testing formats.

**Cautions**

Essay questions do favor the students with better writing skills. Reliability is low because of inconsistencies in scoring, and scoring is a slow process. Scoring problems occur when the instructor hasn’t constructed firm grading criteria before starting the reading/grading process.

**Scoring essay tests**

In the interests of fairness and time management, students should be given the scoring criteria to use while writing the exam.

---

The difference between failure and success is doing a thing nearly right and doing a thing exactly right.

—Edward Simmons
How to Get a “A” in Question and Test Selection and Creation

Include a scoring statement such as the statements below for each essay question:

To receive an “A” grade the response for this question would, at a minimum, include:

To receive a “B” grade the response for this question would, at a minimum, include:

Etc.

Also, if you plan to reduce the test score for incorrect spelling and grammar errors, state on the test exactly how the test grade will be adjusted.

What’s an extended essay question?
There are two ways to construct an essay question. First, the extended-response type. This type is written using language that allows a wide-open range of student answers.

An example of this type is:

Q.) “Describe the role of the Environmental Protection Agency in controlling hazardous wastes.”

What’s a restricted essay question?
The other option, the restricted-response type adds specificity.

For example, this type of question might read as follows:

Q.) State three reasons why you agree and three reasons why you disagree with current EPA policy in the oil and gas industry and their drilling activities in coastal plains.

The Interpretive Exercise
The Interpretive Exercise (IE) is the least recognized question type among the popular choices. It’s a high flexible question format, that can be used to construct questions requiring higher-order thinking skills to answer.

When creating an IE, the stem is expanded from a single sentence statement or question to a paragraph. The paragraph could be a narrative setting up the parameters of a real life situation containing an array of facts, procedures, relationships, etc.

The student is then asked in a series of true-false or multiple-choice questions to interpret and solve problems concerning the data given in the stem.

The Interpretive Exercise (IE) question stem can be created from or contain almost any information source: photos, drawings, maps, cartoons, x-rays, graphs, tables, data sets, poems, and short stories, etc.
How to Get a “A” in Question and Test Selection and Creation

Test takers are asked to preform a wide variety of application, analysis and synthesis tasks and then answer a wide variety of questions. For example, after reading a description of two businesses that includes some financial data in chart form, students can be asked to make predictions, calculate quantities, efficiencies or various ratios and then make judgements about which business has the highest profit potential in the upcoming six months.

Science, health and technology instructors can use the IE to test students’ knowledge of designs, laboratory tests and troubleshooting procedures. Social science educators can create organizational, societal or individual case studies and then ask a variety of questions to engage students with the data.

For example, in a technology class exam an IE question might contain a flow chart of a process, for instance, a flow chart of a heating and cooling system for a commercial building complete with computer data readout at key points.

Then, the students can be asked a variety of questions about what it means if the data points are out of tolerance. Or they can be asked to state corrective actions they should take.

IEs make excellent test questions for group tests because of the wide possibility for interpretation and discussion among group members.

Which question type is the best?
The more restrictive the language the easier the question will be to score, but the less insight you will have into the student’s ability to organize and express knowledge. These are some of the trade-offs to consider in designing an essay question.

It has become popular with some instructors to offer the students a selection of optional questions. Before you do this, you should consider the effect of these questions on the reliability of the test. By offering options you tend to lose the common basis for measuring accomplishment that occurs when all students answer the same questions.

Grading tips
Essay grading is much easier when you have developed an answer outline for each question. Include in the outline the content and organization of the information you expect to see in the answer along with the number of points to be awarded for each section. This scoring aid is best prepared during test construction.

—George W. Crane
How to Get a “A” in Question and Test Selection and Creation

Decide whether you will grade by the point method or the rating method. The point method involves comparing the answer to your key and assigning the points earned for each question. The rating method is based on creating piles of paper with similar answer quality.

For the rating method, five-to-ten categories are the most often cited number of quality piles used. The answers are then read and their characteristics are rated. Grades are then assigned according to quality category. Remember that it is best to read all answers to each question before grading the next one; also, it is easier to be more objective if you don’t know the author’s name.

Your turn: End of Section 5. Review Questions

Check Your Test Question Construction Knowledge

1.) T. F. To minimize ambiguity, the best way to construct a multiple-choice question stem is to write a complete sentence.

2.) A restricted-type essay question contains which of the following question components?
   1. Distracters.
   3. Amplification responses.

3.) Which question type usually consists of two columns of information and a task of finding relationships?
   1. IE.
   2. MC.

4.) What is an important first principle of test writing?
   1. Add humor whenever possible.
   2. Each question should be related to a learning objective.
   3. Always add “All of the above” to the distracter list.
   4. The stem should be a phrase or sentence fragment.

5.) A good way to gain insight into test-taker thinking is to add which feature to the question?
   1. A true-false format.
   2. A witty remark to stimulate thinking.
   3. Use “None of the above” as a choice.
   4. Add an answer justification response block.

6.) How should a fill-in-the-blank question be written?
   1. Write a statement or question with at least three blanks.
   2. Limit the question or statement to one blank.
   3. Use “None of the above” as a choice.
   4. Add a scoring statement concerning grammar and spelling.

Click here to check your answers.
Section 6. Course Test Planning

For the benefit of your students and yourself, develop a brief but flexible Course Testing Strategy during the planning stages of your course each term. Then publish this plan in your syllabus to communicate your evaluation and testing strategy to your class members.

A “TFS Course Testing Strategy Grid” is provided to help you with this task. Your Strategic Testing Plan will result in the pleasant feeling of having exerted control over this complex teaching area, and you will not be pushed into the unenviable position of having to make last-minute decisions the night before a test must be given. Your students will very much appreciate knowing how you plan to evaluate and grade their academic progress. When they know the plan up front they can make much more informed “stay” or “drop” decisions during the first weeks of class.

The first strategic planning decision you should make is to decide when and how often to test. Some tests, such as final tests, are scheduled by your institution. If this is the case at the institution where you teach, you have little choice about when you can administer a final, but unit tests, weekly quizzes and daily classroom learning assessments can be scheduled as you see fit.

Although there are no precise guidelines for estimating an optimum testing frequency, generally, the more you test the more reliable will be your basis for determining a grade, and the better you will be able to target the precise learning you want to measure. Don’t rule out a daily one-question classroom learning assessment. These can be as simple as, “write the most important thing you learned in this class session on a 3 x 5 card and turn it in.”

Also, the scheduling of more evaluations rather than fewer allows students to see earlier in the learning process how they are doing so that they can correct their learning problems before it’s too late. With “one midterm and a final” testing strategy, there is little students can do to improve their performance in time to make a difference. To see the value of using a planning grid, review the sample provided.

There is no doubt about the role of questioning in learning. Questions drive learning and provide vital learning reviews. Tests give the learner a chance to identify weak areas and points of confusion. No test is perfect. It’s better to test more often than to get too tied up in the pursuit of perfection and offer your students too few tests, quizzes and performance checks. Ask, ask, ask and your students will learn more and to a deeper level than you may have thought possible, and you’ll be teaching for success.
The TFS Course Testing Planning Grid is used to plan an effective evaluation strategy. Mr. Tesla plans to combine traditional and modern approaches by scheduling in-class quizzes, laboratory proficiency checks and a final with a take-home midterm and end-of-chapter student summaries instead of in-class chapter tests.
How to Get a “A” in Question and Test Selection and Creation

What to test?
When planning a course, deciding what to test is a crucial decision. Valid tests cover only the course’s learning objectives or topic list. If you have written learning objectives or are following a list of objectives or topic outlines provided to you, you have a ready-made checklist for test construction.

Remember, your evaluation strategy should be to carefully devise a revealing sample of the whole possible universe of learning achievement contained in your course. Use tests and quizzes to track achievement and to identify learning problems your students are experiencing. This diagnostic information can help you steadily improve your teaching and your course. And the same information can help your students learn to be better learners.

Planning strategy tips
Plan your course evaluation strategy both as formative, (in process checks) and as a summative, (end of course) final examination (if desired or required by your institution). As in any sampling process, increasing the frequency and variety of methods should increase the accuracy of the results.

An important planning note: Plan with a positive attitude. Testing is a serious endeavor, and playing power games with tests trivialize the student’s efforts to learn and achieve. Never use a test as a punitive measure for poor learning achievement or inappropriate classroom behavior. If you are angry with your class for any reason it may be wise to put off test planning until the situation is corrected.

When and where to test?
Teaching philosophies today offer more evaluation choices than ever before. Since class time is strictly limited, the more tests, quizzes and other evaluation activities you schedule the less time your students will have for new learning activities. Some faculty see tests as valuable learning activities, therefore, the time allocated for testing is not lost. Others see diverting learning time to testing as a major concern. Take-home tests can be a solution to the time crunch.

Take-home tests?
One answer to the time problem is to employ several take-home tests to take the place of some in-class evaluations. Jacobs and Chase, Developing and Using Test Effectively, recommend that to make take-home tests successful in a competitive learning environment, it’s essential that you set time- or page-length limits on the answers to be submitted. Otherwise, some students will likely skew the results with test overkill.

Group testing
As an additional option, student group- and pair-completed tests can be completed either in class or as take-home tests. This type of testing is known as Collaborative...
How to Get a “A” in Question and Test Selection and Creation

testing. Students turn in individual score or answer sheets but work together to arrive at the answers.

Some research has shown that better students do not tend to carry unprepared students as some instructors often fear or surmise. The key to good collaborative testing is the composition of the study teams. If they are functioning productively, collaborative testing should go well, too. As you can see, you have a great many options to include in your course testing plan including critiques of artworks and writing journals.

Also, your college or university may offer computer-managed testing services in a computer laboratory setting. With adaptive-testing programs you can develop objective tests that consist of a selection of different questions each time the test is taken. This type of system eases the chore of providing make-up tests.

Finally, during your test-planning process give thought to how you will handle make up-tests. Even with a major test scheduled, it’s rare that 100 percent of the class will make the test for one reason or another. You need to have a policy to deal with this situation, especially a final test.

Important considerations

Here are several questions that need special consideration while planning your testing strategy:

- Do you record an incomplete if a student misses a final or major test?
- Do you reduce the final grade by one or two levels if the final is missed?
- If a test can be made up what test will you give?
- Is it fair to give the same version of a test as a make-up?
- Could alternative assignments or projects be developed that would adequate test learning achievement in the content area?

Congratulations!! You have completed the TFS Quick-Study on testing and evaluation. The Teaching For Success planning charts, tables and checklists included in this Quick-Study can save you time and energy in designing, creating, scoring and evaluating your tests. When you have mastered testing and evaluation skills you have mastered the fifth Critical Success Factor of good teaching.

Be sure to browse through the reference section when looking for good resources by which to further increase your skills. Look for SuperIdeas and QuickTips on testing in the monthly editions of Teaching For Success.

Knowledge and skills that you develop in this area will make you a better instructor. Your students will benefit from more accurate and valid assessments of their learning. You will save valuable instructional time and have the satisfaction knowing that your testing and evaluation strategy is based on sound principles.

Lastly, we have provided a blank TFS Test Planning and Analysis Grid in the Job Aides section for you to print as needed. Feel free to use and modify these job aides.
How to Get a “A” in Question and Test Selection and Creation

Section 7. References

The TFS recommended list of references will help you learn more about testing theory and test construction:


As an instructor concerned with the success of your students and your own professional development, your goal is to teach students how to succeed. You should help to facilitate their learning and study skills enough to allow them to view tests as ways to demonstrate what they have learned—not as anxiety-producing instruments to be endured.

Now that you have completed this basic QuickStudy on testing, plan to keep your testing and evaluation skills growing by reading the SuperIdeas and QuickTips available in Teaching For Success, the monthly faculty development and training letter. If your college or university does not provide you with TFS, ask your Faculty Development key person or chief instructional administrator to subscribe.

Other TFS QuickStudys available: Positive Classroom Discipline, Making Small Groups Work, Make Better Presentations, Constructing an A+ Syllabus and Planning a Successful Lesson. Call 800-757-1183 or click on the following URL, teachingforsuccess.com for more details. Send feedback on this QC to jack@teachingforsuccess.com.

It’s up to you. Specific testing decisions all come down to your best judgment of your intended purpose, your students’ abilities, and your own willingness to give time and energy, not only to creating the test, but also to scoring and grading the evaluations that you choose.
Section 8. Job Aides

The TFS Test Planning and Analysis Grid (TPAG)

Course:
Type, Basis, Purpose and Scope:

<table>
<thead>
<tr>
<th>Performance Levels</th>
<th>Knowledge Target (%)</th>
<th>Comprehension Target (%)</th>
<th>Application Target (%)</th>
<th>Analysis Target (%)</th>
<th>Synthesis Target (%)</th>
<th>Evaluation Target (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning Objectives</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#Q; Type</td>
<td>Q Focus</td>
<td>#Q; Type</td>
<td>Q Focus</td>
<td>#Q; Type</td>
<td>Q Focus</td>
<td>#Q; Type</td>
</tr>
<tr>
<td>Quantity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quantity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quantity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quantity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quantity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### How to Get an “A” in Question and Test Selection and Creation

<table>
<thead>
<tr>
<th>Study Sections</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Click to Go)</td>
</tr>
<tr>
<td>Overview</td>
</tr>
<tr>
<td>Testing Basics</td>
</tr>
<tr>
<td>Testing Talk</td>
</tr>
<tr>
<td>Test Planning</td>
</tr>
<tr>
<td>Question Construction</td>
</tr>
<tr>
<td>Course Test Planning</td>
</tr>
<tr>
<td>References</td>
</tr>
<tr>
<td>Job Aides</td>
</tr>
<tr>
<td>Answers</td>
</tr>
<tr>
<td>Contact Information</td>
</tr>
</tbody>
</table>

#### Table for Question and Test Selection

<table>
<thead>
<tr>
<th>Total Questions</th>
<th>Percentage of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**A TFS Quick Study**
## How to Get a “A” in Question and Test Selection and Creation

### TFS Course Testing Planning Grid

<table>
<thead>
<tr>
<th>Course:</th>
<th>Instructor:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>What to Test</strong></td>
<td><strong>Where and When</strong></td>
</tr>
<tr>
<td>Unit or Chapter</td>
<td>Learning Objectives or Topics</td>
</tr>
<tr>
<td>Ch</td>
<td></td>
</tr>
<tr>
<td>Ch</td>
<td></td>
</tr>
<tr>
<td>Ch</td>
<td></td>
</tr>
</tbody>
</table>
# How to Get a “A” in Question and Test Selection and Creation

| Unit or Chapter | Learning Objectives or Topics | Date | In class | Take Home | Quiz | Unit Test | Mid-term | Final | Per | Critique | Chap- | Journ |
|-----------------|-------------------------------|------|----------|-----------|------|-----------|----------|-------|-----|----------| Ter | al |
| Ch              |                               |      |          |           |      |           |          |       |     |          |     |     |
| Ch              |                               |      |          |           |      |           |          |       |     |          |     |     |
| Ch              |                               |      |          |           |      |           |          |       |     |          |     |     |

**Study Sections**

(Click to Go)
- Overview
- Testing Basics
- Testing Talk
- Test Planning
- Question Construction
- Course Test Planning
- References
- Job Aides
- Answers
- Contact Information

**References**

**Job Aides**

**Answers**

**Contact Information**
How to Get a “A” in Question and Test Selection and Creation

Learning Level Question Construction Chart

<table>
<thead>
<tr>
<th>Learning Level</th>
<th>Initial Term in Question Statement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge</td>
<td>Describe, identify, label, select, name.</td>
</tr>
<tr>
<td>Comprehension</td>
<td>Predict, estimate, distinguish, summarize.</td>
</tr>
<tr>
<td>Application</td>
<td>Calculate, solve, show, use, change.</td>
</tr>
<tr>
<td>Synthesis</td>
<td>Categorize, organize, rearrange, combine.</td>
</tr>
<tr>
<td>Evaluation</td>
<td>Justify, interpret, compare, contrast.</td>
</tr>
</tbody>
</table>

Do your questions pass the test?

<table>
<thead>
<tr>
<th>Q Type</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>MC</td>
<td>MC and True-false are highly flexible and especially work well to test the lower learning levels such as knowledge, comprehension and application. Four-choice questions are the best compromise. Can be machine scored.</td>
<td>While MC questions can be written to test Synthesis, Analysis and Evaluation levels, these are difficult to write. An essay or Short-answer question may be a better choice for testing at the higher learning levels.</td>
</tr>
<tr>
<td>True-false</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAT</td>
<td>Matching allows you to quickly test a range of knowledge with fewer questions—great, for example, in testing terminology.</td>
<td>More time consuming to create than multiple-choice and not well suited for Analysis, Synthesis, and Evaluation levels.</td>
</tr>
<tr>
<td>FB and SA</td>
<td>Fill-in or Short-answer questions are very easy to construct. Great for quizzes and are better indications of knowledge retention because the correct answer is not visible.</td>
<td>Usually not machine compatible. Compatible with all learning levels. Easy to construct, time consuming to grade with large classes.</td>
</tr>
<tr>
<td>ES</td>
<td>Student selects facts, organizes information; draws conclusions, uses higher learning skills such as analysis and synthesis. Can be used in open-book and take-home and group formats.</td>
<td>Scoring criteria must be created and added to question. Time consuming to grade accurately and fairly—favors the students with better writing skills.</td>
</tr>
<tr>
<td>IE</td>
<td>Compatible with Application, Synthesis, Analysis and Evaluation levels. Tests complex areas.</td>
<td>Requires more time effort to construct good questions than essay, but it’s easier to score.</td>
</tr>
</tbody>
</table>
Section 9. Answers

Answers: End of Section 2. Testing Basics
Check your Testing Basics

1). Why is a list of learning objectives or learning outcomes essential to have before a test is created? Write your answer in the field below.

1). A: Each test question must be tied to a learning objective.

2). Name the three parts of a formal learning objective statement?

2). A: Part 1, Give; Part 2, Action or will do statement; Part 3, Performance standard or how-well statement.

Click to go on to Section 3. Testing Talk.

Answers: End of Section 3. Testing Talk
Check Your Test Talk

1.) Which testing-talk term describes a systematic process of making value judgements?

The correct term is evaluation.

2.) A final exam is an example of which type of evaluation?

1. Formative.
2. Summative.
3. Diagnostic.

3.) A test that measures exactly what it was designed to measure has what characteristic?

1. Subjectivity.
2. Objectivity.
3. Reliability.
4. Validity.

4.) A question that features responses that can be rapidly and accurately graded by anyone is an example of a:

1. Subjective question.
2. Objective question.
3. Performance question.
4. Valid question.

5.) A teacher grading a test finds that the average score is 78 percent. The instructor then looks for groups of scores higher and lower than the average and assigns these groups “A”, “B”, “D” and “F”, with a “C” being from 80-76 percent. This is an example of an instructor using which kind of referenced test?

1. Norm-referenced.
2. Objective-referenced.
3. Criterion-referenced.
6.) A test that yields a similar range of scores class after class would have what level of reliability?
   1. **High.**
   2. Medium.
   3. Low.

   **Click to go on to Section 4. Test Planning**

**Answers: End of Section 4. Test Planning**

1.) Step one of test planning includes which of the following activities?
   1. **Define the purpose of the test.**
   2. Assign a Performance level to each question.
   3. Track the number of questions created.
   4. Analyze the question type percentages.

2.) The second performance level of Bloom’s Taxonomy of Cognitive Learning is described by which of the following terms?
   1. Formative.
   2. Analysis.
   3. **Comprehension.**
   4. Application.

3.) What information is entered in the cells forming the left most column of the TPAG?
   1. Question type.
   2. **Chapter and associated learning objectives.**
   3. Analysis figures of the question usage percentage.
   4. Number of multiple-choice questions used.

4.) What was the main content focus imbalance revealed by the TPAG in the test planning sample?
   1. Chapter one may be over emphasized.
   2. **Fifty percent of the questions cover Chapter three.**
   3. Too many true-false questions are used.
   4. The total number of test questions should be increased.

5.) What was the Performance Level imbalance revealed by the TPAG in the test planning sample?
   1. Too many Knowledge level questions asked.
   2. Too many Application level questions asked.
   3. Too few Application level questions asked.
   4. Too many Synthesis questions asked.

6.) What is the recommended way to estimate the approximate length of a test that you are creating?
   1. Fill out a TPAG
   2. Give the test to colleague to take.
   3. Use two-minutes per question as a good rule-of-thumb.
   4. **Fill out a Test Time Analysis Grid.**

7.) You notice that the top students in the class finish a test in one-half the time required by the slowest students. What is indicated about the test length?
   1. It’s too long.
   2. **It’s about right.**
   3. It’s too short.
Click to go on to Section 5. Question Construction

Answers: End of Section 5. Question Construction

1.) **T.** F. To minimize ambiguity, the best way to construct a multiple-choice question stem is to write a complete sentence.

2.) A restricted-type essay question contains which of the following question components?
   1. Distracters.
   3. Amplification responses.
   4. **Response-limit statement.**

3.) Which question type usually consists of two columns of information and a task of finding relationships?
   1. IE.
   2. **MC.**
   3. **MAT.**
   4. ES.

4.) What is an important first principle of test writing?
   1. Add humor whenever possible.
   2. **Each question should be related to a learning objective.**
   3. Always add, All of the above” to the distracter list.
   4. The stem should be a phrase or sentence fragment.

5.) A good way to gain insight into test-taker thinking is to add which feature to the question?
   1. A true-false format.
   2. A witty remark to stimulate thinking.
   3. Use, None of the above” as a choice.
   4. **Add an answer justification response block.**

6.) How should a fill-in the blank question be written?
   1. Write a statement or question with at least three blanks.
   2. **Limit the question or statement to one blank.**
   3. Use, None of the above” as a choice.
   4. Add scoring statement concerning grammar and spelling.

Click to go to Section 6. Course Test Planning

--- End of Study ---

We wish you much success in your evaluation efforts!

Go to teachingForSuccess.com for more Quick Study Titles

Contact Information

QuickStudys™ created by Teaching For Success® and Pentronics Publishing, 1270 Mt. Rainier Dr., South Lake Tahoe, CA 96150. Jack H. Shrawder, Publisher and Editor. Contact: jack@teachingforsuccess.com or 800-757-1183.

Your comments and improvement suggestions are always welcome.