

**Writing and Reviewing Assessment Items:  
Guidelines and Tips**

**by John Painter, Ph.D.**

**September, 2004**

**Family and Children's Resource Program  
Jordan Institute for Families  
School of Social Work  
University of North Carolina at Chapel Hill**

## Writing and Reviewing Assessment Items

The focus of this document is on the practical aspects of planning and writing test items. A well-designed test should represent the instructional objectives of the curriculum; assessment items should be clearly written and fair. There are three basic steps to test construction: planning, development, and review.

<b>1. Step 1: Planning .....</b>	<b>3</b>
<b>2. Step 2: Development .....</b>	<b>6</b>
<b>3. Step 3: Guidelines for review of test items .....</b>	<b>9</b>

### **Step 1: Planning. Specifying What is to be Assessed**

- 1. Clarify instructional objectives.** Decide what it is students should learn. Include all the objectives that you think are important. Remember that broad, general objectives are of little value in developing test items.
- 2. Define the purpose of the test.** The types of items developed for a test will depend on the purpose of the test. Three common types of test are:
  - a. *Pre-instructional assessments.* Tend to cover a broad range of topics and skills, and does not attempt to make fine discriminations about individual performance.
  - b. *Interim mastery tests.* Administered before moving from one instructional unit to another. These tests tend to be relatively brief, but composed of highly specific questions designed to sample a range of cognitive objectives in proportion to what is important for final outcomes.
  - c. *Mastery tests.* Used for assigning grades, deciding on promotion, or selection for special recognition. These tests are administered infrequently, but carry disproportionate importance because the scores can have long-term implications. It is very important that these tests are thoroughly scrutinized and validated.
- 3. Specify what will be tested and at what level.** Any given objective can be tested at many different levels. Decide whether a student's ability to recall facts, apply principles, or demonstrate mastery is the expected outcome.
- 4. Develop a test blueprint.** After deciding upon important objectives and the specifics of what will be tested develop a plan to guide the number and difficulty of test item construction. See Table 2 for an example.

**Table 1: Summary of Bloom’s Taxonomy of Educational Objectives for the Cognitive Domain. Ordered from most to least frequently used.**

<b>Level</b>	<b>Definition</b>	<b>Question Cues:</b>
<b>Knowledge</b>	<ul style="list-style-type: none"> <li>• Observation and recall of information</li> <li>• Knowledge of dates, events, places</li> <li>• Knowledge of major ideas</li> <li>• Mastery of subject matter</li> </ul>	List, define, tell, describe, identify, show, label, collect, examine, tabulate, quote, name, who, when, where, etc.
<b>Comprehension</b>	<ul style="list-style-type: none"> <li>• Understanding information</li> <li>• Grasp meaning</li> <li>• Translate knowledge into new context</li> <li>• Interpret facts, compare, contrast</li> <li>• Order, group, infer causes</li> <li>• Predict consequences</li> </ul>	Summarize, describe, interpret, contrast, predict, associate, distinguish, estimate, differentiate, discuss, extend
<b>Application</b>	<ul style="list-style-type: none"> <li>• Use information</li> <li>• Use methods, concepts, theories in new situations</li> <li>• Solve problems using required skills or knowledge</li> </ul>	Apply, demonstrate, calculate, complete, illustrate, show, solve, examine, modify, relate, change, classify, experiment, discover
<b>Analysis</b>	<ul style="list-style-type: none"> <li>• Seeing patterns</li> <li>• Organization of parts</li> <li>• Recognition of hidden meanings</li> <li>• Identification of components</li> </ul>	Analyze, separate, order, explain, connect, classify, arrange, divide, compare, select, explain, infer
<b>Synthesis</b>	<ul style="list-style-type: none"> <li>• Use old ideas to create new ones</li> <li>• Generalize from given facts</li> <li>• Relate knowledge from several areas</li> <li>• Predict, draw conclusions</li> </ul>	Combine, integrate, modify, rearrange, substitute, plan, create, design, invent, what if?, compose, formulate, prepare, generalize, rewrite
<b>Evaluation</b>	<ul style="list-style-type: none"> <li>• Compare and discriminate between ideas</li> <li>• Assess value of theories, presentations</li> <li>• Make choices based on reasoned argument</li> <li>• Verify value of evidence</li> <li>• Recognize subjectivity</li> </ul>	Assess, decide, rank, grade, test, measure, recommend, convince, select, judge, explain, discriminate, support, conclude, compare, summarize

**Planning: Example of a Test Blueprint**

Here is an example of a test blueprint for a select competencies from the curriculum

**Table 2: Test Blueprint**

**Foster Home Licensing**

<b>Competency</b>	<b>Knowledge</b>	<b>Comprehension and Application</b>	<b>Analysis, Synthesis, and Evaluation</b>	<b>Total Number of Items</b>
The worker understands the historical, philosophical, and legal basis of child welfare practice.	1	1		2
The worker knows the individual and family characteristics and dynamics that contribute to effective foster and adoptive parenting.	2	2	2	6
The worker knows the values of family-centered practice, including family preservation, permanence for children, preservation of parent’s and children’s rights, client self-determination, reasonable efforts, and respect for individual and cultural differences.	2	3	2	7
	<b>5</b>	<b>6</b>	<b>4</b>	<b>15</b>

Decisions embodied via a test blueprint:

- How many items are to be constructed for a specific Competency?
- What cognitive level are questions going to address?
- The Statewide Partnership has specified that the number of items per competency be determined by the amount of time spent on that competency during training.

## Step 2: Development: Basic Tenants of Test Item Writing

### Maxims of Item Writing in General

1. Keep the reading level difficulty of test items appropriate in relation to the individuals who are taking the test, unless of course the purpose is to measure verbal and reading abilities. You do not want language difficulties to interfere with a student's opportunity to show what she knows.
2. Write items as simply as possible, making sure that students know exactly what information is being requested. Irrelevant details, grammatically or logically sloppy construction, or unnecessarily sophisticated vocabulary can contribute to confusion. Often the writer has such a clear idea of what is being solicited that she does not realize that there are alternative interpretations to the question posed. For example:
  - Poor: The United States Constitution was written in \_\_\_\_\_.
  - Better: The United States Constitution was written in the year \_\_\_\_\_.
3. Test for important ideas, information and skills—not trivial detail. It is often easiest to write questions about specific details, but is that information you want to test? While knowledge about specific facts is important on the way to mastering higher-level cognitive skills, it is usually not necessary to test each detail on the path to that higher-level skill.
4. Avoid using direct quotations or verbatim statements from a test or article to form a question. It is not unusual to save time by constructing a test item from a passage in a book or article. Taken out of context such material can be ambiguous or confusing, and it also encourages students to focus primarily on memorization rather than understanding. However, it is acceptable to present a specific passage then refer to it in subsequent questions.
5. If an item is based on opinion or authority, indicate Whose opinion or What authority so that the item clearly has only one correct or best answer.
6. Avoid using interrelated items. If getting the correct answer to one item depends on correctly answering a previous item then undue weight is given to the first item. It is ok to refer to the same stimulus material, as long as getting the correct answer to one item does not depend on the correct answer to another item.
7. Avoid irrelevant clues and 'give-away' questions. Unintentional clues include a tendency to reserve / avoid a certain position for the correct answer (e.g., response D is usually correct / incorrect) or having subject-verb agree between stem and correct response only. Also avoid use of words such as *never*, and *always*.

## Writing and Reviewing Assessment Items

8. *Avoid trick questions.* A sure way to alienate students is to develop questions that confuse trivial detail with major issues. Consider the following True-False question:  
T     F     John L. Kennedy was the youngest person to serve as President of the United States.  
The answer is technically false because President Kennedy's middle initial is F. Even if the student recognizes the difference she would not know if it is a typographical error. The student should not be expected to assess the test writer's motivation or typing skill when answering a question.
9. *Ask questions about single ideas; avoid double-barreled questions.*
10. *Have someone else review your items.*

### **Writing Multi-Choice Items**

A Multiple-choice question consists of a statement or question (the stem), followed by a number of possible responses (the options, distractors, or alternatives).

#### **Terms:**

**Multiple-Choice Item:** This is the most common objective-type item. The multiple-choice item is a test question which has a number of alternative choices from which the examinee is to select the correct answer. It is generally recommended that one use 4 or 5 choices per question, whenever possible. Using fewer alternatives often results in items with inferior characteristics. The item choices are typically identified on the test copy by the letters A through E.

**Stem:** This is the part of the item in which the problem is stated for the examinee. It can be a question, a set of directions or a statement with an embedded blank.

**Options/Alternatives:** These are the choices given for the item.

**Key:** This is the correct choice for the item.

**Distractors:** These are the incorrect choices for the item.

#### **Difficulty:**

The difficulty level of a multiple choice item will depend on two things:

- A. The thought process called for by the stem. This can range from recalling factual information to evaluating new information.
- B. Similarity of the options. Items with options with a high degree of conceptual similarity require greater understanding in order to identify a correct response.

## Writing and Reviewing Assessment Items

1. Distractors should portray a single concept that appears plausible to students who have not mastered the material. Here are some guidelines for writing good distractors:
  - a. Base distractors on the most frequent errors made during class discussions or on assignments related to the concept.
  - b. Use words in the options that are associated with words in the stem.
  - c. Use concepts from the instructional material that have similar vocabulary or were used in the same context as the correct answer.
  - d. Use distractors that are similar in content or form to the correct answer
  - e. Make the distractors similar to the correct response in terms of complexity, sentence structure, and length.
  
2. Most multiple-choice items should have three to five options. Although more options will reduce the likelihood of guessing the correct answer, the quality of the distractors is most important.
  
3. Place repetitious words in the stem, instead of in each option.
  
4. Use 'None of the above' and 'All of the above' sparingly.
  
5. In general the stem of an item should be meaningful without having to read all the options. In other words, the stem should be a complete sentence or idea.

*Poor:* The President of the United States

  - a. is chosen by the electoral college
  - b. can serve a maximum of two four-year terms

*Better:* The President of the United States is chosen by the

  - a. electoral college
  - b. members of the Senate
  
6. Avoid elaborate explanations or including 'background information' in an effort to make an item clearer.
  
7. Options should come at the end of the statement, not in the middle. Although this is not always possible, items written this way are usually easier to understand.

*Poor:* The city of \_\_\_\_\_ is the capital of California.

*Better:* The capital of California is \_\_\_\_\_ .

### Step 3: Guidelines for Review of Test Items

The following guidelines are recommended for reviewing individual test items. When you review an item, write your comments on a copy of the item indicating your suggested changes. If you believe an item is not worth retaining, suggest it be deleted.

1. Consider the **item as a whole** and whether
  - a. it measures knowledge or a skill component which is worthwhile and appropriate for the examinees who will be tested;
  - b. there is a markedly better way to test what this item tests;
  - c. it is of the appropriate level of difficulty for the examinees who will be tested.
  
2. Consider the **stem** and whether it
  - a. presents a clearly defined problem or task to the examinee;
  - b. contains unnecessary information;
  - c. could be worded more simply, clearly or concisely.
  
3. Consider the **alternatives** and whether
  - a. they are parallel in structure;
  - b. they fit logically and grammatically with the stem;
  - c. they could be worded more simply, clearly or concisely;
  - d. any are so inclusive that they logically eliminate another more restricted option from being a possible answer.
  
4. Consider the **key** and whether it
  - a. is the best answer among the set of options for the item;
  - b. actually answers the question posed in the stem;
  - c. is too obvious relative to the other alternatives (i.e., should be shortened, lengthened, given greater numbers of details, made less concrete).
  
5. Consider the **distractors** and whether
  - a. there is any way you could justify one or more as an acceptable correct answer;
  - b. they are plausible enough to be attractive to examinees who are misinformed or ill-prepared;
  - c. any one calls attention to the key (e.g., no distractor should merely state the reverse of the key or resemble the key very closely unless another pair of choices is similarly parallel or involves opposites).

## Writing and Reviewing Assessment Items

### References

Bloom, B.S. (Ed.) (1956). *Taxonomy of educational objectives: The classification of educational goals: Handbook I, cognitive domain*. New York ; Toronto: Longman, Green.

Krathwold, D. (1993). *Methods of Educational and Social Science Research*. New York: Longman.

Thorndike, R. and Hagen, E. (1955). *Measurement and Evaluation in Psychology and Education*. New York: Chapman & Hall, Limited.

Worthen, B., Borg, W., and White, K. (1993) *Measurement and Evaluation in the Schools*. New York: Longman.

Wollack, James A. and Cohen, Allan S.. *Handbook on Test Development: Helpful Tips for Creating Reliable and Valid Classroom Tests*. Testing & Evaluation Services University of Wisconsin-Madison available on line at:  
<<http://wiscinfo.doit.wisc.edu/exams/Handbook%20on%20Test%20Construction.pdf>>