# Table of Contents

The Importance of Documentation.................................................................................................................. 3
Purpose ......................................................................................................................................................... 3
Basic Development ......................................................................................................................................... 4
Writing Test Items ......................................................................................................................................... 5
  Objective vs. Constructed Response............................................................................................................... 6
    Objective Tests ........................................................................................................................................... 6
    Constructed Response Tests ....................................................................................................................... 7
Test Your Test.................................................................................................................................................. 9
The purpose of this guide is to provide a brief overview of the most important elements of Cognitive Test Construction. It is intended as a quick reference source to help you maintain a systematic method for developing valid and reliable evaluation instruments.

The Importance of Documentation

Documentation of each step of test development is critical. Keep in mind that documented plans are your only defense if the test is subject to internal and/or external scrutiny. (Shrock, S., Coscarelli, W. 2000)

Purpose

Just as you would analyze and identify the need for an instruction intervention, you must also be clear about the purpose of the test. Ask yourself what the test is meant to accomplish. The following list gives brief descriptions of typical test purposes for use in corporate training:

- Prerequisite Tests – ascertain that learners have skill levels required for mastery of course objectives.
- Entry Tests – identify skills or range of skills that learners already possess to possibly modify course delivery
- Diagnostic Tests – used during or immediately following instruction to assess mastery of objectives. May help identify problem areas.
- Posttests – administered after instruction to assess learner mastery of objectives
- Equivalency Tests – used to determine if learner has already mastered objectives and may bypass instruction
- Certification – May exist independent of instruction. Often used by corporations or NGOs to recognize advanced skills or knowledge

(Shrock, S., Coscarelli, W. 2000)
Basic Development

1. Review course learning objectives for clarity. Revise if necessary. Remember that your goal is to construct test items that will provide valid and reliable assessment of the stated learning/performance objectives.

2. Match questions to knowledge level required by objective.

3. Only use questions about salient points of the covered topic.

4. Draft a large number of test items (at least 4-5 on each objective)

5. Select well-designed and best coverage of topic items.

6. Sequence the items logically. They should typically appear in the order of their application to the task.

7. Begin with easier items to avoid discouragement.

8. Prepare instructions to learners.

9. Prepare scoring key.

10. Before implementation, have items reviewed by SME for face and content validity.

11. After several pilot uses, analyze for reliability and validity.

12. Reliability is increased with higher numbers of questions. To maintain interest, vary types of items.

13. Use clear language.

14. Questions should not include clues to the answer.

15. Question wording should not give clues to any other item answers.

16. Avoid use of negatives.

17. Phrase questions to test learner's ability to apply knowledge. Use paraphrase. Change context.

18. Vary placement of correct answers.

19. Avoid trick questions.

20. For closed-ended items, write questions with goal of only one correct answer.

21. Do not make answer to any question dependent on answering a previous question correctly.

(Newby, A. C., 1992), (Shrock, S., Coscarelli, W. 2000)
Bloom's Taxonomy – Objectives Written at What Level?

- Knowledge level – definitions, procedures, formulas, poems, etc.

- Comprehension level – restate in own words, translate from one form to another, apply designated rules, recognize previously unseen examples of concepts

- Application level – decide what rules are pertinent and implement rules to solve problem

- Analysis level – Break complex situations into component parts and figure out how parts relate to and influence one another.

- Synthesis level – create totally original material. Closed-ended questions cannot measure attainment of synthesis-level questions.

- Evaluation level – judge appropriateness of some object, plan, design for some purpose. Cannot be assessed with closed-ended questions.

(Shrock, S., Coscarelli, W. 2000)
Writing Test Items

Objective vs. Constructed Response

The relationship between learning objectives and test items forces some choice between Objective and Constructed Response items because Objective test items are generally only suitable for assessing lower learning levels. Therefore, written tests intended for the Synthesis and Evaluation levels will necessarily include Constructed Response items.

Objective Tests

The major types of Objective Test items are:

- Multiple Choice
- True/False
- Matching

Multiple Choice – The advantages of Multiple Choice items are that they can be constructed to measure attainment of learning objectives through the Analysis level and they are easy to grade. Disadvantages of Multiple Choice items are that they measure recognition rather than recall and they are fairly difficult to construct. They are also susceptible to guessing of correct answers.

Multiple Choice items consist of a stem, which is the incomplete statement, and the response items, which include the correct answer and several distractors. Learners must choose the correct answer from several alternatives.

Guidelines for Writing Multiple Choice Items:

1. Use three to five distractors. More than this may be too challenging for the question writer and may overwhelm the reader.
2. Construct test items to measure learning objectives. Avoid trivia. Use credible distractors.
3. The stem should include as many words as possible to reduce the size of the response alternatives.
4. Use clear language and write the stem as a complete thought.
5. The correct answer should not stand out because of length, phrasing or grammatical structure.
6. Avoid use of limiting modifiers in the distractors (e.g., always, never)
7. Avoid indiscriminate use of “all of the above” and “none of the above”
8. Construct stem statements as positives. If the use of negatives is unavoidable, use graphic emphasis (bold, underline, caps) of “NOT” to reduce confusion.

True/False Items - True/False Items require the learner to decide if a statement is correct or incorrect. They are good measures of recognition of factual information. The advantages of true/False Items are that they are relatively easy to write and very easy to score. The disadvantages are that they are
Quick Guide for Written Cognitive Test Instruments

limited to assessing only the knowledge and comprehension levels, there is a high probability of guessing the correct answer and they test recognition rather than recall.

Guidelines for Writing True/False Items

1. Statements must be written unambiguously. Avoid use of limiting (always, never) or ambiguous (sometimes, maybe) modifiers.
2. Limit the statement to a single thought or idea.
3. Include equal number of true and false items.
4. Use True/False items sparingly for assessment.

Matching Items - In a Matching Items test, the learner identifies a relationship between lists of entries (Descriptors) in one column with a list of responses (Options) in a second column. Matching tests are generally used to test recognition. They work best when each listing forms a category of related items.

The advantages of Matching Items tests are that a large amount of material can be condensed into a smaller space than would be required for Multiple Choice items, and chances of guessing correct answers are less than with other objective test items. The disadvantage is that Matching Items are limited to assessing only the knowledge level of learning.

Guidelines for Writing Matching Items

1. Limit the number of Descriptors to six or seven.
2. Make the items as short as possible – a word, phrase or brief sentence
3. Options should be able to be used more than once to reduce chances of guessing.
4. Options should include several distractors.
5. All test items require clear instructions to learners. Special emphasis on instructions is required for Matching test items.

Constructed Response Tests

Constructed-Response test items require learners to plan and express answers in their own words. In written tests, Construction-Response items must be included for assessment of Synthesis and Evaluation learning levels. The major types of Constructed-Response tests are:

- Completion Items
- Short Essay and Long Essay

Completion – Completion items require the learner to fill in a word or words in response to an incomplete statement of a question. Advantages of Completion items are that a large amount of content can be tested in a limited time period and they test recall rather than recognition. Disadvantages are that they are limited to assessment of only the first four learning levels and scoring, while easier than other Constructed-Response items, is still subject to unanticipated learner response.
Quick Guide for Written Cognitive Test Instruments

Guidelines for Writing Completion Items

1. The single most important rule in writing Completion items is to construct the question statement so that only one answer is correct.

2. Use only one blank per item.

3. Place the blank near or at end of sentence.

4. Try to use single word answers.

Essay Items - Short Essay items are useful in assessing through the highest levels of learning. Learners are instructed to respond to a question with a limited number of words, sentences or paragraphs or a limited amount of space (number of pages). The limitation requires the learner to provide a focused response. In contrast, Long Essay questions offer learners more opportunity to expand or defend a rationale. All essay questions are challenging to grade, with increasing difficulty according to the amount of expansion or divergence allowed. Advantages of Essay items are that they are relatively easy to construct, they can assess higher-order learning and they can provide deeper insights into learner mastery of concepts. Disadvantages of Essay items are that they require more time and preparation by learners, they penalize poor writers who may be knowledgeable, and grading is subjective and time-consuming.

Guidelines for Writing Essay Questions

1. Write focused and specific questions.

2. Describe the grading criteria in detail; i.e., spelling and grammar, points awarded for each section, etc.

3. Write a model answer.

4. Do not allow learners to choose which essay question to answer.

(Shrock, S., Coscarelli, W. 2000), (Pershing, J., 2004)
Test Your Test

Before implementation, ask the SME (and any other appropriate persons) to review items for Face and Content Validity. Revise as suggested.

Plan to “pre-test” the instrument several times with groups who are demographically similar to the intended learner group. Use results of the “pre-test” for formative evaluation of the instrument.

(Shrock, S., Coscarelli, W. 2000), (Newby, A. C., 1992), (Pershing, J., 2004)
Addendum

A Case for Performance Testing

In some contexts, standards of performance are included as part of the instructional objectives. In these cases, performance testing is the only meaningful assessment of attainment of objectives.

Before constructing a performance test, determine if product, process, or both are to be evaluated. Process testing evaluates the confidence, care and accuracy of the learner’s performance of the procedure. Product testing evaluates the quality and/or quantity of something produced by the learner. Most performance testing includes assessment of both of these elements.

Resources and safety issues are often determining factors in deciding whether performance testing will occur under realistic or simulated conditions.

Performance tests consist of two components:

- Observation Support Tool (rating scale or checklist)
- Observer or Rater

The Observation Support Tool must reflect job-related competencies. It is written to provide adequate coverage of performance objectives. Typical examples of Observation Rating Tools are:

- Checklists – determine that sequential steps of a process are performed. Do not rate quality.
- Rating Scales – may be norm-referenced or criterion-referenced. Scales must be descriptive with clearly defined numeric values
- Rubrics – focus more on quality of the learner’s performance or product, often evaluating related skills in a complex task. Rubrics are especially useful in providing feedback to learners.

Validity of performance tests depends on rating items which are closely matched to stated performance objectives.

Direct observation of performance can be very effective in evaluation. In order to minimize subjective ratings, Observers and Raters must be trained to consistently recognize and identify specific behaviors in accordance with the descriptions of the Performance Support Tool. Observer/Raters should be as unobtrusive as possible and refrain from verbal commentary while observing.

Reliability in performance testing relates to consistent correlation of Observer/Rater assessments. It is important to use several Observer/Raters and to compare their assessments to determine reliability.

(Shrock, S., Coscarelli, W. 2000), (Newby, A. C., 1992)
Quick Guide for Written Cognitive Test Instruments

References

