

Learning Outcomes Assessment: Developing and Refining Criterion- Referenced Tests

Gail R. Ryser, PhD
Director

Testing, Research-Support,
and Evaluation Center

Outline of Workshop

1. Review learning outcomes.
2. Define items, tests, and item formats.
3. Learn how to write good exam items.
4. Learn a technique for analyzing exam results that relate to performance criteria and learning outcomes.

SACS Criteria

CS 3.3.1 Institutional Effectiveness

The institution ***identifies expected outcomes for its educational programs*** and its administrative and educational support services;

assesses whether it achieves these outcomes;

and ***provides evidence of improvement based on analysis of those results***

Learning Outcomes

- ✓ Statements that describe what students are expected to know, think, and be able to do **by the time of graduation**.
- ✓ Focus on intended **abilities**, **knowledge**, **values**, and **attitudes** of the student after completion of the program.

Learning Outcomes

- Specific
- Measurable
- Achievable
- 5~8 per program

Example Learning Outcomes

- **Students will be able to communicate information effectively in writing, orally, and graphically.**
- **Students will be able to compare and contrast the major theoretical perspectives in philosophy.**
- **Students will acquire the ability to identify, formulate, and solve engineering problems.**

Direct Assessment Methods

- ✦ Class Assignments (paper, presentation, report...)
- ✦ Capstone Project
- ✦ Performance Project
- ✦ Direct Observation
- ✦ Portfolios
- ✦ External examiner
- ✦ Standardized exam
- ✦ Locally developed exam
- ✦ Certification and licensure exams
- ✦ Simulations

Thorndike

- Constructing good test items is probably the most demanding type of creative writing imaginable. The item writer must understand the content measured by the item and must determine whether the cognitive demand will involve recall, understanding, or application.

Items and Tests

- A test item is the basic unit of observation in any test.
- A test item is intended to measure some aspect of human ability generally related to learning or training.
- A test is a measuring device intended to describe numerically the amount of learning under uniform, standardized conditions.
- Typically a set of items are intended to measure a domain of knowledge or skills or cognitive ability.

Item Development Process

1. Create or use an inventory of items in a test bank.
2. Develop test specifications or a test blueprint (for course outcomes that relate to program learning outcomes).
3. Identify the number of items needed for each learning outcome you plan to cover in your course.
4. Evaluate the performance of items.
5. Place good items in your test bank.

Performance Criteria	Basic Facts	Application	Analysis
Creates a test using test design steps.			
Interprets the results of item analyses.			
Illustrates item formats by writing several types of items.			
Analyzes results of exams to improve instruction.			

Evaluating Item Performance

- Item Difficulty
 - The percentage of examinees who answered the item correctly. Varies from 0.0 to +1.0.
- Item Discrimination
 - The degree to which the item discriminates between the top students and the bottom students. Varies from -1.00 to +1.00.
- Breakdown
 - The frequency of choices chosen by the students. Examine breakdown separately for top and bottom students.

General Item-Writing Guidelines

Content Guidelines

1. Every item should reflect specific content and a single specific cognitive process.
2. Use novel material to measure understanding.
3. Keep the content of an item independent from content of other items on the test.
4. Avoid the following: overspecific or overgeneral content; opinion-based items; trick items.

Style and Format Guidelines

1. Format items vertically instead of horizontally.
2. Edit for clarity, grammar, punctuation, and spelling; proofread.
3. Simplify vocabulary so reading comprehension does not interfere with testing the content.
4. Minimize reading time.

Writing the Stem Guidelines

1. Make the stem as brief as possible.
2. Place the main idea of the item in the stem, not in the choices.
3. Avoid negative words.

Writing the Options Guidelines

1. Use as many options as possible, but three seems to be a natural limit.
2. Vary the location of the right answer according to the number of options-assign the location randomly.
3. Place the options in logical or numerical order.
4. Keep options homogeneous in content, grammatical structure, and length.
5. Avoid using:
 - a. None of the above
 - b. All of the above.
6. Avoid negative words such as not or except.
7. Make all distractors plausible.

Multiple Choice Formats

- Conventional
 - Three parts: stem, correct choice, distractors
 - Stem should provide a complete idea of the knowledge to be indicated in selecting the right answer.
 - Correct choice is the one and only right answer.
 - Distractors are unquestionable the wrong answer. Must be plausible to test takers who have not learned the knowledge or skill the item is supposed to measure.

Conventional Continued

- Variations
 - Question Format
 - Who is John Galt?
 - Best Answer
 - Find the word that most nearly means the same as anarchy.
 - Incomplete Stem
 - John Galt is a character in Ayn Rand's novel "Atlas Shrugged" who is remembered for his _____
 - Incomplete stem may not be best format, especially for limited English proficient individuals.
 - Never use blanks in the middle of the stem or question.

Complex MC Format

- Offers three choices regrouped into four options:
 - A. 1 and 2
 - B. 2 and 3
 - C. 1 and 3
 - D. 1, 2, and 3
- Cons:
 - More difficult for test taker.
 - Test-taking skills have a greater influence.
 - Produces items with lower discrimination and lower score reliability.
 - Requires more reading time and takes up more space.

Matching Format

- Two parts: stems or statements, options
- Additional Guidelines
 1. Provide more options than stems.
 2. Put options in logical order (e.g., numeric order).
 3. Stems should be longer than the options.
 4. Number stems and use letters for options.
 5. Keep content of the options homogeneous.
- Extended Matching
 - Uses a lead-in that is either a scenario or vignette.
 - Places the items in a real life context.

Alternative Choice Format

- Conventional MC with two options.
 1. What is the most effective way to motivate a student?
 - a. Intermittent praise
 - b. Consistent praise
- Pro
 - Only one distractor must be written.
- Con
 - Probability of guessing the correct answer is 50%.

True False Format

- Pros
 - Easy to write and score.
 - Reduction of reading time.
- Cons
 - Tendency to test recall of trivial knowledge.
 - Large error component due to guessing.
 - Tend to be less reliable than other types of format.
- Additional Guidelines
 1. Balance the number of TF statements.
 2. Use simple declarative sentences.
 3. Write items in pairs (but only include one on a single test).
 4. Trivial details should not make a statement false.

Multiple True/False Format

- Uses a lead-in statement , followed by a series of TF statements.
- Pros
 - Produces higher reliability when compared to conventional MC.
 - Efficient in terms of item development and examinee reading time.
- Cons
 - The variety of content appears to be limited to testing understanding of examples and nonexamples, characteristics and noncharacteristics.
 - Difficult to keep items independent (i.e., one item cues another item).

Context-Dependent Item Sets

- Two parts
 - Introductory stimulus (e.g., scenario, vignette)
 - 2 to 12 items (test sets) related to the stimulus
- Good for measuring reading comprehension, problem solving, pictorial, interlinear
- Guidelines for writing test sets
 1. All items should be on single page.
 2. Use any format that seems suitable (e.g., conventional MC, matching, MTF).

Exercise

1. Identify which learning outcome (if any) and course objectives the two multiple choice questions you brought would assess.
2. Partner with a neighbor and critique each others MC questions.
3. Based on the critique and our discussion, rewrite your MC questions.

Exam analysis

- Identify test items that match course objectives and learning outcomes.
- Calculate overall student performance for each item.
- Calculate the average performance for items assessing same objective.
- Determine the level of competence.

Exam Analysis

(Reference: Mi-Suk Shim, PhD; Division of Instructional Innovation and Assessment; The University of Texas at Austin)

Objective	Exam 1	Exam 2	Assignment 1	Assignment 2	Assignment 3
Analyze communication situations and audiences	1 <u>89%</u> 5 <u>82%</u> Avg <u>86%</u>				
Write business documents that are grammatically correct and use appropriate business style					
Deliver effective business presentations					
Develop effective interpersonal communication skills					
Use communication technology appropriately and effectively	3 <u>75%</u> 11 <u>85%</u> 15 <u>45%</u> Avg <u>68%</u>				
Conduct research and use it to complete written and oral reports					

Further Assistance

- Dr. Beth Wuest
Director, Academic Development and Assessment
Office: JCK 150
Email: wuest@txstate.edu
- Dr. Gail Ryser
Director, TREC
Office: Lower Commons, room 100
Email: gr16@txstate.edu

(Many of the guidelines found in this presentation are from:
Haladyna, T. M. (2004). *Developing and validating multiple-choice test items*.
Mahwah, NJ: Lawrence Erlbaum.)