

MODULE 12: EVALUATION TECHNIQUES

Cognitive goals

At the completion of this module the student-instructor should be able to:

- 12.1 Define evaluation, formative evaluation and summative evaluation
- 12.2 Distinguish between formal and informal evaluation
- 12.3 Identify various types of evaluation, and the advantages and disadvantages of each
- 12.4 Understand specific types of test items and the advantages and disadvantages of each
- 12.5 Understand general guidelines for test construction
- 12.6 Define reliability
- 12.7 Define content validity
- 12.8 List several examination resources

Psychomotor goals

At the completion of this module, the student-instructor should be able to:

- 12.1 Develop two examples of correctly constructed test items for cognitive evaluation in each of the following categories:
 - Multiple choice
 - True/false
 - Matching
 - Completion
 - Essay
- 12.2 Develop a skills checklist to evaluate a psychomotor skills performance

Affective goals

At the completion of this module, the student-instructor should be able to:

- 12.1 Explain why evaluation is important to the total quality control of an EMS program
- 12.2 Explain why formative and summative evaluations are both important in the evaluation process

Declarative

- I. Why this module is important
 - A. Without a mechanism to evaluate the student you will never know if you have achieved the objectives and goals of instruction
 - B. The evaluation process helps determine strengths and weaknesses of students and your program
 1. Program evaluations help improve the quality of the instruction
 2. Student evaluations help determine whether students are progressing satisfactorily
 3. Evaluations can also be used to determine if an individual is compatible with the EMS field (by targeting the affective domain)
 - a. See the appendix for a sample affective domain evaluation tool

- C. Entry level EMS instructors may not design and develop test items (questions), but should have familiarity with the concepts of evaluation and how to construct solid test items
 - 1. Even if using a prepared test bank, you should understand how to determine if these questions are well written and match the objectives of your lesson plan

- II. Evaluation
 - A. Process of making a value judgment based on information from one or more sources
 - B. A mechanism of determining student progress toward, or the attainment of, stated cognitive, psychomotor, and affective objectives
 - C. The evaluation process should look at two components
 - 1. Instruction as provided by the instructor
 - 2. The performance of the student on course and lesson objectives

- III. Purpose of evaluation
 - A. Provide feedback to student on progress or performance
 - B. Provide student gratification and motivation to succeed
 - C. Measure effectiveness of teaching style and content of lesson
 - D. Measure effectiveness of the educational program in meeting written goals and objectives

- IV. Formative evaluation
 - A. Ongoing evaluation of the students and instruction conducted throughout the course
 - 1. Compare the overall goal of instruction, lesson objectives and the content to the performance by the students
 - 2. Compare the objectives of the course to the testing strategy
 - a. Cognitive component: testing knowledge
 - b. Psychomotor component: testing skill performance
 - c. Affective component: testing attitudes, beliefs, ethics and morals
 - B. Formative evaluation is important in gaining insight early in the program
 - 1. Use this information to make changes in the program, to provide remediation, or to redirect presentations
 - C. Methods of performing formative evaluation during a course or class
 - 1. Module or section testing within a larger topic area is a form of formative evaluation
 - 2. "Taking their temperature" is an informal method of obtaining a quick response of student's questions or to clarify content just delivered
 - a. Two of the many methods to take their temperature
 - i. One minute paper: ask student to write their response to a question then compare their answer to another student's
 - ii. The muddiest point: ask students to write any questions they have on note cards and collect prior to a break
 - a.) After the break begin with a review of the most common questions
 - 3. Give frequent, short-duration written or practical drills or quizzes
 - a. The intent is to provide feedback to both the student and instructor on the progress of the student

- V. Summative evaluation

- A. Summative evaluation is performed at the completion of the delivery of a larger section of material, a module or program
 - B. Provide feedback to the students of their successful mastery of the content
 - C. Determine the effectiveness of teaching strategy and to improve future teaching performance
 - D. Methods of performing summative evaluation during a course or class
 - 1. Survey tools
 - a. Gather opinions about various aspects of the course and instruction
 - 2. Comparison of course and program outcomes
 - a. Determine if all goals and objectives were met
 - 3. Final examinations: practical and written
 - 4. Test item validation
 - a. Determine if questions were valid
 - b. Psychometric assessments can be performed to validate tests and questions
 - i. This level of evaluation is beyond an entry level EMS instructor
 - E. A formative evaluation can also be summative
 - 1. Depending upon the context in which it is used, a test may represent formative or summative evaluation
 - a. For example: a multiple-choice final exam given at the end of a topic will be both formative and summative
 - i. It is summative because it represents the end of that topic area
 - ii. It is formative because it represents only a part of a course
- VI. Formal and informal evaluation
- A. Both formal and informal strategies are critical to the success of courses and programs
 - B. Some of the evaluation strategies listed can be conducted formally or informally
 - C. Formal evaluation
 - 1. A structured instrument to assess student's attainment of interim and/or terminal course objectives
 - 2. A formal written examination can determine a grade for a course or serve as a means to continue in the program
 - 3. If you review the test and allow students to challenge questions (prove it wrong through the use of textbooks and class notes, etc.) it can serve as a powerful learning tool
 - 4. Problems with formal evaluation techniques
 - a. Place stress on the student (especially the ill-prepared ones)
 - b. May not provide a mechanism for remediation or retention when they represent a final summation of learning
 - D. Informal evaluation
 - 1. Less structured method of assessing student achievement used primarily to provide corrective feedback to both the student and instructor
 - 2. Informal evaluation tools may not be graded or the instructor may not record the grades
 - 3. Student benefit: identify weakness (and strengths) and offer suggestions for improvement; may serve as a "wake-up" call
 - 4. Instructor benefit: compare results from the class to identify trends and problems and to develop corrective instruction or remediation

5. If the instructor makes an evaluation informal it may cause conflict when students have the expectation of a formal evaluation
 - E. Problems with formal evaluation techniques
 1. May not allow for remediation or retraining
 - F. Problems with informal evaluation techniques
 1. Students may not perceive the value in these instruments because grades are not recorded
 - a. Instructors may not wish to spend class time doing informal evaluations
 - b. Instructors may not grade or provide the student feedback on informal evaluations further diminishing their importance in the student's eyes
- VII. Written examinations
- A. Types
 1. Multiple choice
 2. True/false
 3. Matching
 4. Completion
 5. Essay
 6. Terminal/certifying
 - B. Terminology
 1. Item: common instructional design term for all of the components of a written examination question including the question, correct (or best) answer and incorrect answers
 2. Stem: part of the item that is first offered, it may be written as a question or an incomplete statement, the stem is often called the "question"
 3. Distracter: an answer to a test question that is a false or incorrect answer designed to be a plausible alternative to the correct answer
 4. Key: the correct (or best) answer to the item
 - C. Source of test items: the course and lesson objectives
 1. Test items should come from the course objectives and lesson plan
 - a. If you are testing information you have not directly covered in the class you should ensure students have been directed to this information via reading assignments, projects, or some other form of independent study
 - D. Advantages of written examinations
 1. Can be used with a large number of students
 2. Measures cognitive objectives
 3. Provides for consistent scoring
 4. Grading and compiling result is quicker than for other types of examinations
 - E. Disadvantages of written examinations
 1. Time consuming to develop
 - a. Difficult to develop adequate measurements for the higher order levels of the domains of learning
 2. Complex validation procedures
 3. Could discriminate against students with reading difficulties
 - a. Poorly written items may evaluate a student's reading ability more than they evaluate knowledge of the material
 4. Cannot measure skills performance

- a. Questions can be asked about the procedure to perform, but the actual skill demonstration cannot be evaluated via this type of test
- VIII. General guidelines for written test item construction
- A. Test must be related to objectives and developed from a blueprint
 1. A blueprint is an outline for the test
 - a. Include test items on each objective
 - b. Decide the depth and breadth (level) to cover for each item
 - B. Exam must be an appropriate type
 1. Considering your domain of learning and how far into the domain (high or low level or level 1, 2 or 3) you want to go
 - a. This can help determine the appropriate instrument to use
 2. The following are a partial list of suggestions to follow
 - a. Low level cognitive: multiple choice, matching, true/false, simple completion (fill-in-the-blank) or short answer essay, and oral exam
 - b. High level cognitive: long and short answer essay, fill-in-the-blank, some true/false and completion, oral exams, projects (case studies for example), and observational reports
 - c. Low level psychomotor: rote skills, oral, and observational reports
 - d. High level psychomotor: situational scenarios, projects (designing scenarios for example), and observational reports
 - e. Low level affective: oral, short-answer essay, projects (opinion papers for example), and observational reports
 - f. High level affective: oral and situational scenarios, projects (group designed presentations for example), and observational reports
 - C. Organize exam in a logical manner
 1. Group like items (similar content area) together on a written exam
 2. Have questions follow a linear or logical sequence in an oral or situational (scenario based) examination
 - D. If timed, allow an appropriate amount of time to answer questions or perform a skill
 1. Determine if timing the test is appropriate
 - a. When preparing for a timed licensure or certification exam
 - b. Mirror timing strategy of terminal exam in your preparatory exam
 - c. If the goal is for students to think or act quickly
 2. Some suggested timing strategies
 - a. One minute per item for a standard multiple-choice test
 - b. Allow 2-3 minutes to read a scenario then one minute for each multiple choice question
 - i. Allow longer time to read a scenario and compose an essay answer
 - c. Allow longer time to respond to a situational skill than a rote one
 - d. Allow longer response time for a higher level question (in any domain) than a lower level one
 - E. Provide clear complete directions about the test
 1. Can or cannot write on the test
 2. Use a pencil to fill in the answer sheet
 3. Time limit
 4. Whether or not breaks will be allowed during the test

- F. Have another instructor review the examination for clarity and completeness
 - 1. Be sure exam is legible, free of typographical, grammatical, spelling and content errors
- G. Be consistent in the design strategies you are using for the graphics (fonts) on a written examination
 - 1. Use all capitals or all lower case letters consistently throughout the test for both the key and distractors, both in the numbering strategy (a, b, c, d or A, B, C, D) and for the first word of the key and distracter
 - 2. Be consistent in the use of punctuation
 - 3. Use a consistent strategy to draw attention to material in the test
 - a. Be consistent with the use of underline, bold or italics
 - 4. Position key and distractors appropriately
 - a. Observe for answers that build logically
 - i. If your answer choices are the letters "a," "b," "c" and "d," place them in that order
 - ii. Place number answers in ascending or descending order
- IX. Specific types of written test items: multiple choice
 - A. Common method of conducting formal and informal evaluation
 - B. Written national and state certification and licensing examinations are multiple choice
- X. Limitations of multiple choice questions
 - A. Bias cueing occurs from leading students to the correct answer by the way the stem is worded or from the grammar choices
 - B. Inadequate stems require students read all of the answer choices before selecting an answer
 - C. Negatively worded stems should be avoided
 - 1. Students are used to looking for positively worded stems and can be tricked by (or misread) negative ones
 - D. Questions should not build on previous questions information
 - 1. Exception is in testing the sequencing of steps
 - E. Avoid questions written with a fill-in-the-blank segment in the middle of the stem
 - 1. Difficult to read and the meaning may be skewed
 - F. Avoid “all of the above” or “none of the above” as a distracter
 - 1. Recognition of one other incorrect distractor immediately eliminates “all of the above” as a possibility
 - 2. Recognition of a couple distractors as correct (or possibly correct) leads the student to guess that “all of the above” is the correct answer
 - 3. “None of the above” can be an alternative if the question is a mathematical (computational) one
 - G. When you need to combine answers to obtain a correct answer all of the possible combinations should be used to make up the distracters
 - 1. Questions with only four options result in over fifteen answer combinations making this impractical
 - 2. Overlapping responses are a problem
 - a. If the question asks for a range and one answer offers a single number it can be immediately eliminated

- b. If distracters overlap into the range you are looking for it can be confusing to the student
 - i. Example: the correct range is 6-8 and the choices are 2-6, 4-7, 6-7 and 5-9
- XI. Specific types of written test items: true/false
- A. Includes a complete statement and a two choice alternative of true or false
 - B. Limitations
 - 1. Item must be limited to the two choices of true or false
 - a. Does not allow for any gray area
 - 2. Difficult to construct good items in a positive voice
 - a. Avoid negatively worded statements using "is not", etc.
 - 3. Avoid extreme answers that include the absolute statements "always" or "never"
- XII. Specific types of written test items: matching
- A. Typically two columns of information are offered with the intent of selecting items from one column and matching them to items in the other column to form correct or complete statements
 - B. Limitations
 - 1. Works best with definitions and terms or with simple concepts with obvious relationships
 - 2. Difficult to properly design
 - a. Multiple matches may be possible within the columns
 - 3. Items used must bear some similarity
 - a. Unless you were matching terms with definitions it is useless to match terms like humerus, beta blocker, and inferior because the answers would be obvious
 - 4. Unclear directions how matching will occur
 - a. Explain if students will use each term 1 time or multiple times
 - b. Explain if single or multiple answers are needed to complete a match
- XIII. Specific types of written test items: completion
- A. Fill-in-the-blank
 - B. Statements with part of their information omitted so student must complete the statement
 - 1. Limitations
 - a. Enough information must be included for the student to glean the intent of the statement without leading the student to the answer
 - b. Meaning may be unclear and several answers may emerge as correct
 - c. The answer space may provide a problem
 - i. Gives a hint to the student if a blank line is used for each word of the answer
 - ii. A single line may mislead student to think the answer is a single word when it is actually two or three words
- XIV. Specific types of written test items: essay
- A. Short answer
 - 1. Requires a bulleted list of responses or several questions to complete
 - B. Long answer
 - 1. Requires students to provide a lengthy prose style answer

- C. Limitations of both types
 - 1. May not be effective for measuring the lower levels of the domains of learning
 - 2. Time consuming and sometimes difficult to grade
 - a. Grading can be very subjective
 - i. Group grading is an alternative
 - b. Scoring can be difficult as you try and assign a point value to the various components of the expected response
 - i. Rubrics are helpful tools for grading essay questions because a rubric will describe the criteria for each grade level
 - a.) Example: For an "A" the student must provide all the correct information and write in complete sentences without committing any spelling errors and for a "B" the student must provide 80% of the required answer and commit one to three spelling errors
 - 3. Students often write illegibly because of time pressure or may try to add information at the end making the response difficult to follow
 - 4. Some students thought processes do not flow easily a linear progression causing an unfair disadvantage in a timed test
 - 5. Students may include much more information than desired in an attempt to be thorough
 - 6. If students do not understand the question, they may provide a very well thought out, but incorrect, answer
 - a. Should you award partial credit for a well-constructed incorrect answer?
- XV. Specific types of written test items: terminal/certifying
 - A. Final summative examination with the intent of granting permission to attempt a licensing or certifying examination
 - 1. Most often a multiple-choice examination
 - B. Examination given with the intent of granting a certification or license by a regulating body such as a state or the National Registry of EMTs
 - 1. Requires successful completion of one or a combination of two or three of these examinations to obtain certification or licensure
 - a. Written examination (generally multiple choice)
 - b. Oral examination
 - c. Practical skills examination
- XVI. Post written examination quality review by students
 - A. Will students be allowed to retain the test?
 - 1. Advantages
 - a. Provides a learning aid for later testing
 - b. Provides examples of your style of question writing
 - 2. Disadvantages
 - a. Generally eliminates the test (and maybe all of the questions) from reuse
 - B. Allow students the opportunity to review the test
 - 1. Highlights areas of weakness for further study and remediation
 - 2. Highlights areas of weakness in the presentation of the material
 - 3. Can help control bias or discrimination concerns when students see what other students missed

4. Promotes a climate of fairness when students can challenge questions, answers, or the wording of a question
 5. Can be used as a learning aid for both the student and instructor
- XVII. Post written examination quality review by faculty
- A. Compile the results, including an accounting of incorrect answers
 - B. If the upper one-third of the group missed a specific item determine the following:
 1. Is the test item keyed correctly?
 2. Is the test item constructed properly?
 3. Is it free from bias, confusion and errors in grammar and spelling etc.?
 4. Was the content covered in class?
 - a. If not, were the students directed to it through self-study?
 - C. If the lower one-third of group missed a specific item:
 1. Which distracters were most attractive? (in other words, were most often selected)
 - a. Improve distracters that were not attractive
 - b. Consider a distracter well written if it is not selected by the upper- third of the class but it is selected by the lower third
- XVIII. Oral examinations
- A. Exams in which both questions/instructions and answers are given out loud by a student to an instructor, or group of instructors
 - B. Advantages
 1. Evaluates “quick thinking” or reactions
 2. Evaluates the student’s thought processes
 3. Can be evaluated by multiple listeners simultaneously
 - C. Disadvantages
 1. Limited number of students may be examined at any one time
 2. Difficult to standardize
 3. Examiner may unintentionally give clues to the examinee
 4. Time-consuming and labor-intensive
 5. Subjective
 6. To be fairly administered, a great deal of attention and concentration is required on the part of both the evaluator and the student
 - a. Unexpected distractions can impact upon the test
 - b. Instructors may be required to evaluate a large number of candidates with little opportunity for breaks
 - i. Leading to uneven evaluations over time
 - ii. May also lead to identification of themes or trends with unfair emphasis or focus on those repeated mistakes, i.e., holding successive students accountable for preceding students performances
- XIX. Project assignments
- A. Advantages
 1. Allows independent completion
 - a. May be done during class or outside of class
 2. Evaluates ability to synthesize data
 3. Individual projects for specific learning styles or preferences

- a. Students may select from a group of recommended projects or encouraged to develop their own
 - b. Can promote autonomy and independent learning
 - 4. Allows students to work in groups
 - a. They can develop people skills and conflict resolution skills
 - b. Students can learn from each other and stronger students may tutor weaker students
 - B. Disadvantages
 - 1. Difficult to standardize
 - 2. Possible plagiarism
 - 3. If not carefully designed, they may measure only the product excluding the process
 - a. Sometimes the process used to produce the product is just as important as the final project
 - i. For example, learning how to find resources to use in solving a problem or developing critical thinking skills
 - b. If a presentation is required, the grade should not weigh solely on the presentation but should also include the content of the presentation
 - 4. When group members are not working together it places unfair workloads on the members who are contributing
- XX. Observational reports
- A. Advantages
 - 1. Can be used for psychomotor or affective evaluation
 - 2. Reliability is inherent due to repeated observation
 - a. Reliability can be increased by increasing the number of observations
 - B. Disadvantages
 - 1. Presence of evaluator may influence student performance
 - a. Student performs better when being directly observed
 - b. Instructor/evaluator may misdirect student resulting in the need for retraining
 - 2. Time-consuming and labor-intensive
 - a. Frequently a one-on-one experience
 - 3. Developing criteria can be a complex task
 - 4. Experiences may not be available at the time they are required
 - a. Student attends a clinical setting and there are no patients
- XXI. Practical examinations
- A. Two basic types: situational and rote
 - 1. Situational: demonstration of a skill in the context of a scenario allowing for manipulation of the outcome or procedure by the student
 - a. Good for evaluating critical thinking skills as well as skills performance
 - 2. Rote: demonstration of the steps of performing a skill independent of manipulation of outcomes
 - a. Generally follows very specific order of steps
 - B. Advantages
 - 1. Most closely approximates actual job performance
 - 2. Allows observation and evaluation of related behaviors and attitudes

3. Allows evaluation of psychomotor skills, decision-making abilities and leadership skills
- C. Disadvantages
1. Difficult to standardize
 2. Time-consuming and labor-intensive to prepare and deliver
 3. Limited number of students may be examined at one time
 4. Instructor providing feedback needs to be clear about expected outcome, whether a situational or rote response is required, and should evaluate accordingly
- XXII. Practical skills evaluation
- A. Rote mechanical skill
1. Requires simple task analysis
 2. Is the easiest skill examination to administer
 3. May or may not reflect the actual field performance capabilities of the student
 - a. Isolated skill performed without "real world" stresses may not adequately evaluate affective and psychomotor domains
- B. Situational skills
1. Evaluates judgment and/or decision-making
 2. Required more elaborate simulations
 3. More difficult to develop and deliver
 4. Is a more accurate predictor of field performance because it asks to student to critically think through a scenario that does not always have an obvious answer
- XXIII. Simple skill evaluation
- A. Define the skill
- B. Determine the degree of expected proficiency
- C. Select a representative sampling if all skills in a given area are not evaluated
- D. Create a written task analysis of the skill if one does not already exist
- E. Develop checklist commensurate with the task analysis
1. Each step should contain some measurable criteria so all evaluators can agree on criteria of successful completion of each step
 2. Look for established standards (i.e. National Registry Practical Skills Examination Sheets) for guidance
 3. Keep the number of steps to a minimum to reduce errors in evaluation
 - a. Allows the evaluator to observe the task as it is performed and complete the evaluation form afterward
- XXIV. Performance evaluations
- A. Determine and define expected outcome
1. Are skills performance or decisions-making process more important in the situation?
 2. How stressful or complicated a situation is the student prepared to handle?
- B. Determine what standards will be used to evaluate the performance
1. Design the situation to be representative of the desired outcome: realistic environment, realistic patient complaints and injuries, manikins vs. real people
- C. Evaluate the resources needed for testing
1. The higher the domain level the more realistic the scenario should be
 2. Simulate situation and responses as accurately as possible

- D. List all activities which should be completed in the situation
 - 1. Prioritize activities and list them in their linear (start-to-finish) sequence
 - 2. Weigh most important aspects and critical criteria appropriately
 - 3. Checklist should contain
 - a. The minimum number of properly ordered steps necessary to complete the task
 - b. Steps which are independently observable and measurable
 - c. An outcome consensus understood by each evaluator
 - 4. Avoid qualification of student performance by the evaluator
 - a. During the examination the evaluator should be free to observe the activity and quantify the behavior (check if it was performed or not) and should not be focused on measuring how “much” they performed each step
 - 5. Assure adequate organization to ensure outcome of a situational-oriented performance evaluation
 - a. Provide a skeletal framework for the evaluator to follow
 - b. If scenario involves patient care include information so instructor can provide consistent feedback to each student
 - i. Example: vital signs for appropriate and inappropriate treatment

XXV. Characteristics of skill/performance evaluations

- A. Objectively measures the performance
 - 1. By the instrument
 - 2. By the observer
- B. Replicability
 - 1. Does the instrument measure similar performances consistently
 - a. From one student to another?
 - b. From one class to another?
 - c. From one location (situation) to another?
- C. Fair standards
 - 1. Standards are known by student and faculty
 - 2. Practice with similar instrument during the training session
- D. Realism
 - 1. Situations, scenarios, and patient information are plausible
 - 2. Reactions or changes in the patients are realistic given the intervention and treatment
 - 3. External distractions are realistic
 - 4. Stress is similar to work environment

XXVI. Reliability

- A. The extent to which an exam is consistent in measuring student performance
- B. Does it measure a behavior or body of knowledge consistently on different occasions?
- C. Does the environment influence consistency?
- D. Do different administrators influence results?
- E. Does it discriminate against groups or individuals?

XXVII. Content validity

- A. The extent to which an examination is representative of a defined body of knowledge; the ability of an examination process to measure the knowledge and skills it was intended to measure, in accordance with curriculum objectives
- B. Are the sub-tests weighted and distributed properly?
 - 1. Do they place an over importance on a single test?
 - a. Is that your intent?
- C. Does it cover a reasonable sample of the knowledge and skill objectives?
- D. Is it an accurate predictor of field performance?

XXVIII. Resources for examinations

- A. Peer instructors within your organization may be a good source for examinations of all types
- B. Formalized instruments from certification and licensing bodies (which may also be validated instruments)
- C. Jurisdiction or state EMS office
- D. National Registry of EMTs for sample multiple choice items, practical skills, and oral examinations
- E. Written examination resources
 - 1. NREMT
 - 2. Publishers test banks
 - 3. EMS textbooks
 - 4. EMS textbook instructor guides
 - 5. Textbooks of practice certification examinations
 - 6. On-line and computer based practice certification tests
 - 7. EMS Internet sites
- F. Practical examination resources
 - 1. NREMT
 - 2. NAEMSE
 - 3. EMS Internet sites
 - 4. EMS textbooks
 - 5. EMS continuing education programs
- G. Oral examination resources
 - 1. NREMT
 - 2. NAEMSE
 - 3. EMS Internet sites
- H. Project assignments
 - 1. NAEMSE
 - 2. EMS textbook instructor guides
 - 3. College or university resources
 - 4. Learning styles / preferences information with practical application suggestions

Bibliographical References

American Psychological Association. (n.d.). *Standards for education and psychological tests*. (revised ed.). Washington D.C.: American Psychological Association.

- Burba, A.C. (1998). Multiple Choice Item Writing. *Domain³*, Winter 1998.
- Johnson, D. W., & Johnson, R. T. (1996). *Meaningful and manageable assessment through cooperative learning*. Edina: Interactive Book Company.
- Judd, R. L. (1998) The Pedagogue's Column The Matter of Advising on Test Construction. *Domain³*, Winter 1998.
- Merwin, S. (1992). *Evaluation: 10 significant ways for measuring and improving training impact*. Resources for Organizations, Inc.
- Waagen, A. (1997). Essentials for evaluation. *ASTD*, Issue 9705.