30. Distractor. In testing, incorrect answers to multiple choice or matching type test items.

31. Domains of learning: the categories or classifications of learning.

a Cognitive - intellectual knowledge and skills.

b. Affective - feelings, attitudes, and values.

c. Psychomotor - physical, hands-on skills.

32. Distinguished Honor Graduate. The student who, as a minimum: attains the highest end-of-course percentile average of 95% or higher; does not score below 80% of possible points on any initial written exam; successfully passes the APFT for record on the first try; successfully passes all practical exams on the initial attempt; successfully passes all FTX performance tasks; does not have any disciplinary actions on record; does not have any adverse documentation on student counseling forms.

33. Enabling Learning Objective (ELO). An objective that supports the terminal learning objective. It must be learned or accomplished to learn or accomplish the terminal learning objective. It consists of an action, condition (unless it is a cognitive unaided ELO), and standard. A TLO does NOT have to have an ELO, but it usually has more than one ELO if it does have ELOs. ELOs may be tested directly (individually) or indirectly (in support of the TLO).

34. End of Course Comprehensive Test (EOCCT). A test administered at the end of the course immediately prior to graduation. EOCCT is designed to measure the student's ability to perform critical tasks taught in the course. It provides feedback on the need for both reinforcement training and course revisions.

35. Entry skills test. A pretest designed to determine if a student already possesses a certain skill or knowledge needed as a prerequisite before undertaking new instruction.

36. Entry test. A test which contains items based on the objective that the intended students must have mastered in order to begin the course.

37. Evaluation. The process of assessing individual and/or unit competency.

38. Evaluation criteria. In training, factors which clearly indicate what an evaluator will accept as proof that a unit or soldier can, in fact, perform the task to standard.

39. Examination. The formal evaluation of a student's achievement of specified learning objectives. The three categories of formal evaluations are: (1) hardware performance examination, (2) non-hardware performance examination, and (3) written examination.

40. Fidelity. In job performance measures, the extent to which actions, conditions, cues, and standards of a job performance measure approximate those of a task. In training devices or simulators, the accuracy with which simulators reflect that which they simulate.

41. Field test. A tryout of any training course on a representative sample of the target population to gather data on the effectiveness of instruction in regard to error rates, criterion test performance, and time to complete the course.

42. Formative Evaluation. A within-course evaluation of student competence. Evaluation administered periodically throughout the block of instruction. This type of evaluation provides feedback to the instructor as to the effectiveness of instruction and identifies areas requiring immediate remediation.

43. Go/No Go. The criteria whereby students cannot partially pass. They either pass (go: meet the standard) or fail (no go: do NOT meet the standard).

44. Hands-on. Student practice or testing on actual equipment, simulators, or training aids.

45. Honor Graduate. The student(s) who, as a minimum: attains an end-of-course percentile average of at least 90%; does not score below 70% of possible points on any initial exam; meets other criteria listed in DHG requirements.

46. Instrument. In testing and evaluation, a test or measuring device that is used to determine achievement (go/no go) or the relative standing of an individual or group. Tests, rating forms, and standard interviews are all evaluation instruments.

47. Item analysis. The process of determining whether a test item is functioning as intended. Alternatively, the use of results on individual test items to determine effectiveness of the item. It can be used to obtain feedback on training deficiencies, score exceptions, and improve future versions of the test.

48. Job fidelity. The degree to which a testing situation truthfully and accurately reflects the job situation.

49. Job performance test. An instrument used to determine whether or how well an individual can perform a job.

50. Learning Analysis. A systematic approach for identifying the supporting skills and knowledges for each stated objective that must be acquired before a soldier can demonstrate mastery of the objectives themselves. It also includes identification of types of performance, selection of methods and media, and selection of training site.

51. Learning objective. A statement of what is to be learned in terms of the expected student performances under the specific conditions to accepted standards. Both terminal and enabling objectives are learning objectives.

52. Lesson. A section into which a course of study is divided. It is the level at which training is designed in detail. A lesson is a single continuous session of formal instruction in a subject. A lesson normally provides an opportunity for students to be told about the task they are going to learn (lecture/conference), shown how to perform the task (demonstration), and given the opportunity to practice the task (practical exercise). The student is also provided with feedback (evaluation) following the lesson.

53. Mark sense form. Computer readable sheets on which the soldier records identifying information and answers to questions.

54. Minimum passing score. The score on an examination or in a course which best separates performers from nonperformers. The passing score on the examination.

55. Norm Referenced Test (NRT). A test that grades a student by comparing his/her performance with the performance of other students. It contrasts with criterion-referenced testing where a student is graded based on a pre-specified performance standard. A NRT indicates who has performed the best or the poorest regardless of whether the best performer is competent; i.e., selection of Distinguished Honor Grad [DHG]) may be a norm-referenced evaluation in courses which do not specify that the DHG must achieve a specific standard, e.g., 90% or above.

56. Objective test. A test whose scoring requires no human judgement.

57. Objectivity. In testing, the degree to which a test is scored the same by two or more scorers acting independently.

58. Percentile. A value on a scale of one hundred that indicates the percent of a distribution that is equal to or below it. For example, if a person scores at the 95th percentile on a test, she/he has done better than 95 out of 100 people who have also taken the test and not as well as the remaining five percent who scored higher than she/he.

59. Performance based instruction. The students learn and develop skills by actual performance. Lesson learning objectives are written with an action verb. Students prove competency by actual performance of the objectives.

60. Performance checklist. The breakdown of a task into elements that must be performed correctly to determine whether each student satisfactorily meets the performance standards described in the objective.

61. Performance measurement. The method or score used to assess the level and quality of behavior on a performance test.

62. Performance-oriented training. Training in which learning is accomplished through performance, or the actual doing of the tasks, under specific conditions until an established standard is met. Hands-on training consisting of a skill demonstrating phase, a skills practice phase, and a skill evaluation phase.

63. Performance test. A test that measures the students' ability to perform a particular task or group of tasks under actual or simulated conditions until an established standard is met. A performance test can be administered using actual hands-on performance, simulated performance, or pencil and paper format (e.g., calculating azimuths, map reading exercise, etc.).

64. Post test. A test administered after the completion of instruction to assess whether a student has mastered the objectives of the course or unit.

65. Power test. A test in which items are usually arranged in order of increasing difficulty and in which examinees are given all the time they need to complete as many items as they possibly can.

66. Practical exercise. A practical application of the actions specified in the lesson objectives which is performed by the student under controlled conditions. Gives the student the opportunity to acquire and practice skills, knowledges, and the behaviors necessary to perform the training objective successfully. There are three categories of practical exercises: (1) Hardware-oriented, (2) Non-hardware oriented, and (3) In-classroom, paper exercises.

67. Pretest. A test administered prior to instruction to determine how much the student already knows and to determine if the student needs to take this particular instruction.

68. Process standard. A standard for a task which consists of a series of steps resulting in the soldier obtaining a single result The task is evaluated by observing the process and by scoring each step or element as it is performed in terms of sequence, completeness, accuracy, or speed.

69 Product standard. A standard for a task which terminates in a

product or outcome which is observable and measurable. The task is evaluated by looking at the product or outcome in terms of completeness, accuracy, tolerance, clarity, error, quality, or quantity.

70. Proficiency. A specific standard of performance which the learner must meet in order to demonstrate mastery of a specified behavior. Attainment of a level of skill and knowledge required to perform a given task to standard under required conditions.

71. Reliability. Reliability is the extent to which a test/test item gives consistent results each time it is used. It is the ability of a test to generate equivalent results under repeated trials. That is, the respondent answers the question in a similar fashion when asked repeatedly.

a. Reliability is a stability measure; it is concerned with the consistency with which test results place students in the same relative position if a test is given repeatedly. In testing, reliability can be defined as the consistency of scores obtained by the same persons (1) when examined with the same test on different occasions; (2) with different sets of equivalent items; or (3) with the same or equivalent items under variable examining conditions.

b. Reliability assesses the degree to which test items/tests are free from errors of measurement; the extent to which the results of different forms of the same instrument are similar. Reliability determines the degree of consistency between two measures of the same thing and the capacity of the test to yield consistent results time after time. It attempts to answer the question, "How consistent are the scores of a given student from one measurement to the next?"

c. Reliability uses statistical means to determine how much error is present. In general, if test results are consistent from one measurement to another, test reliability is high. Test scores assess only a limited sample of student performance at a particular time. Unless the scores can be shown to be reasonably consistent; i.e., generalizable, over different occasions, over different but related test questions, or over different faculty observers, little confidence can be placed in the meaningfulness of the scores. On the other hand, we cannot expect student evaluation results to be free from variation. Numerous factors other than the true skills of a student influence measurement error and the quality of student performance. These factors include:

(1 Student fatigue, emotional strains, temporary fluctuations in

memory, item sampling, or student guessing. These and other individual factors may affect student performance. The goal of developing reliable measures is to minimize these factors (and other variables unrelated to the intent of the measure which may influence the expected outcome).

(2) Extraneous factors such as the complexity of the problems in a mock patient board or the student's familiarity with the specific topic area measured by an examination may influence performance. Extraneous factors such as these introduce a certain error factor in all test scores.

d. Generally, one can increase test reliability - that is, decrease inconsistency in results - by adding more items to a test. Very short tests risk an undersampling of course content; however, very long tests risk learner fatigue.

NOTE: Reliability limits validity. A measure which gives inconsistent results cannot give valid results. Reliability alone does NOT provide assurance of test validity. Without a consistent measurement of each student's performance at different times, it is impossible for test scores to have meaning (i.e., validity).

72. Resident school. A training location other than the soldier's unit where the soldier is a full-time student. Resident schools include noncommissioned officer academies, service schools, training centers, and U.S. Army Reserve schools.

73. Speed test. An instrument in which a time limit is set for completion. The tested individual completes as many items as possible within a given time period. Test items are organized in no particular sequence of difficulty.

74. Standard of performance. A statement which establishes a criteria for how well a task or learning objective must be performed. The standard specifies how well, completely, or accurately a process must be performed or product produced. The standard reflects task requirements on the job or learning requirements in the classroom. If a product is standard, it is stated in terms of accuracy, tolerance, completeness, format, clarity, errors, or quantity. If a process is standard, it is stated in terms of sequence, completeness, accuracy, or speed. Both product and process must be observable and measurable.

75. Stem. A stem is a complete statement or question which poses the problem/asks the question in a multiple choice test item.

76. Student Evaluation Plan (SEP). A method or outline of procedures that will be used to gather data and information for the purpose of assessing the achievement of students enrolled in a course of instruction.

77. Summative evaluation. Evaluation administered at the end of instruction. Represents an overall statement about the effectiveness of instruction; normally used as a basis for assigning grades. An End of Course Comprehensive Test (EOCCT) is an example of a summative evaluation.

78. Sustained poor performance. The point at which a student should be recommended for relief or recycle based on continued below standard performance. Each course director specifies the parameters of sustained poor performance for each proponent course. The parameters are specified within the SEP.

79. Task. A single unit of specific work behavior with clear beginning and ending points and directly observable or otherwise measurable, frequently but not always resulting in a product that can be evaluated for quantity, quality, accuracy, or fitness in the work environment. A task is performed for its own sake; that is, it is not dependent upon other tasks, although it may fall in sequence with other tasks in a duty or job array. A task statement, to be complete, must contain an action verb, an object, and must express the conditions under which the task is performed and the standard which must be met in performance.

80. Task inventory. A listing of all the collective or individual tasks associated with analysis. Also called total task inventory, initial task inventory, or gross task list. Critical tasks for a job are derived from this inventory.

81. Template. In testing, the answer key used to grade tests. It can be a score sheet with correct answers listed or it can be electronic and used with an optical scanner to score tests.

82. Terminal Learning Objective (TLO). The main objective of a lesson. TLOs are the very foundation of the systems approach to developing instruction and to identifying the end product: what the school expects the student to be able to do at the end of training. There is only one TLO per lesson.

a. A TLO is a specific description of the tasks (behavior) the learner is to exhibit after training, the condition(s) under which the performance will take place, and the standard(s) or criterion which must be achieved (exactly what a person must be able to do in order to perform the critical task successfully).

b. A TLO must be developed to cover each critical individual task selected for resident training. The TLO may require exactly the same behavior, condition(s), and standard(s) as the critical task from which it is derived; or the TLO may reflect limitations which do not permit training to the ultimate level of proficiency that is required on the job/in the field. In this case, the TLO is a predictor of future (on the job/in the field) competency.

c. A TLO may stand alone as a translation of a task statement or the TLO may be supported by enabling objectives. A TLO may cover

(1) One critical task: Given a patient with a simulated wound, a field dressing and other necessary supplies, apply a field dressing IAW FM 8-230.

(2) Part of a critical task: "Given a conscious patient and the necessary medical equipment, measure the patient's blood pressure IAW FM 8-230." At a basic level, "measure the patient's blood pressure" may initially be taught as a TLO; however, in subsequent training, "measure the patient's blood pressure" may simply become embedded as part of a more complex critical task (measure a patient's vital signs) and will no longer be expressed as a TLO in succeeding lesson plans.

(3) More than one critical task: Given an unconscious patient and the necessary medical equipment, measure a patient's vital signs IAW FM 8-230. (This TLO covers the critical tasks - measure respiration, pulse, temperature, blood pressure, etc., - which are translated into ELOs in support of the TLO.)

d. When there is a disparity between the TLO and the critical task being taught, it is the TLO standard that the student must achieve to demonstrate competency for course completion. In other words, the standard as taught in the classroom may be different (simulated) than the standard required on the job. (i.e., the instructor will not hit somebody over the head with a baseball bat to demonstrate the task "manage head trauma;" however, the student will be expected to manage a head trauma on the job/in the field, so the task is simulated in the classroom using a learning objective.)

e. All TLOs are subject to testing. Test items must include the same action, condition(s), and standard(s) stated in the TLO.

83. Test. A device or technique used to measure soldier performance on a specific task or subject matter. Also a series of questions, exercises, or other means of measuring the skill, knowledge, intelligence, abilities, or other aptitudes of an individual or group. Also refers to a process by which data is accumulated to serve as a basis for assessing the degree that a system meets, exceeds, or fails to meet the technical or operational properties ascribed to the system.

84. Test compromise. The unauthorized disclosure of a test or test items to prospective examinees which would distort the test results.

85. Test item analysis. A technique used to help spot poor criterion test items. A number of techniques can be used to do this. Acceptable test items discriminate between "masters" and "nonmasters." Unacceptable test items are incapable of making such a discrimination. In item analysis, the analyst looks for test items which are missed by "nonmasters" and passed by "masters."

86. Test item pool. The total set of test items constructed for a specified test, be it a single or multiple objective test. The item pool is reduced by the item analysis and review techniques to yield a final version of the test consisting of the best test items from the pool.

87. Test window. The period of time during which a test may be administered.

88. Testing constraints. Limitations such as time, money, personnel, facilities, and other resources which prohibit job performance measures from being identical to the tasks they measure.

89. Training effectiveness. The determination of whether the training objectives have been met. The evaluator asks whether the unit or soldier meets or exceeds established training standards.

90. Training objective. A statement based on training performance. There are three separate elements which form the structure of the objective: the action which the unit or soldier must be capable of performing; the standard of performance the unit or soldier must meet; and the conditions under which the unit or soldier is expected to perform. Each element expresses a factor essential to the understanding of performance and specifies capability for accomplishing the training objective after completing a specific block of instruction.

91. Unit test. A component of the Army Training and Evaluation Program mission training plan that provides a way to conduct an external evaluation "snapshot" of unit proficiency for the commander.

92. Validity. The degree to which the test measures what it was designed to measure; concerned with what a test measures and how well it does so.

a. The validity of a test is the extent to which the test score(s) allow the instructor to infer (judge, conclude, presume, assume) how well students have attained course objectives.

b. A test by itself is not valid or invalid; a test is valid only if it "works" for its intended purpose. Since a test can be used for different purposes, validation of a test is a situation-specific concept; a test can be valid in one situation and invalid in another. Validity is dependent on the purpose, population, and situational factors in which a test takes place.

c. Validity is a judgement made about the appropriateness of a test (measurement) that results from the scores generated from the test.

d. The two major types of validity that test/training developers

must be concerned with are Content-related Validity and Predictive Validity.

(1 Content Validity.

(a) Determines if the test items actually measure the pre-determined criteria (learning objectives).

(b) Determines if the test items represent the objectives or skills taught and determines the extent to which a test samples course objectives.

(c) Answers the questions: Is what is being tested what was taught? - Are both the instruction and the test(s) used to evaluate learning based on the same behavior (learning objective) specifications?

(2) Predictive Validity

(a) Measures how well the scores on a test (instrument) predict later behavior. Predictive validity is used to estimate future performance based on a present test score.

(b) Estimates the extent to which a test predicts how well students will actually perform on the job. This is accomplished, in part, by comparing test performance with a pre-defined set of criteria (expected job performance).

(c) May also be used to predict current performance: performance on a test/training developer-constructed test may be used to estimate performance on a standardized test.

93. Weighted scoring. Scoring in which the number of points awarded for a correct response is not the same for all items in the test. In some cases, the scoring formula awards more points for one

response to an item than for another.

94. Within course tests. Tests administered during a course of instruction to ensure that all students are achieving the training objectives.

APPENDIX C

AWARDS CRITERIA

NOTE: See also AHS Reg 351-10, Student Academic Awards.

A. Distinguished Honor Graduate. To be nominated as a candidate for DHG, the student must, as a minimum: attain the highest end-of-course percentile average of 95% or higher; not score below 80% of possible points on any initial written exam; successfully pass the APFT for record on the first try; successfully pass all practical exams on the initial attempt; successfully pass all FTX performance tasks; not have any disciplinary actions on record; not have any adverse documentation on student counseling forms. Upon selection as DHG, an official Certificate, prepared by the Academic Records Branch and signed by the Commandant, is awarded to the selected student.

NOTE: The Course Director may add DHG qualification criteria

NOTE: Students who satisfy selection criteria for Distinguished Honor/Honor Graduate will be nominated as candidates. Final selection will be made by the course director. Normally, there will be one DHG selected per class; however, a class may not have a DHG if at least one candidate is not eligible. The course director may determine the number of Honor Graduates per class.

B. Honor Graduate. To be nominated as a candidate for Honor Graduate, the student must: attain an end-of-course percentile average of at least 90%; not score below 70% of possible points on any initial exam; meet other criteria listed in DHG requirements.

Upon selection as Honor Graduate, an official Certificate, prepared by Academic Records Branch and signed by the Commandant, is awarded to the student who achieves the second highest academic average through Phase I, with no first time failure on any exam.

C. Commandant's List. To be eligible for inclusion on the Commandant's List, students must: be in the top 20% (highest point totals) of the number of graduating students in a specific class, including Distinguished Honor Graduate and Honor Graduates; successfully pass the APFT for record on the first try; not have scored below 70% on an initial exam; not have any disciplinary action taken or pending while assigned/attached in a student status; not have any adverse counseling statements on record concerning improper behavior or inappropriate conduct.

D. The Class Leader and Assistant Class Leader may receive letters of appreciation signed by the Course Director.

E. Students are NOT eligible for special recognition if they have been on academic probation or have received disciplinary action.

APPENDIX D

SAMPLE TESTING PROCEDURES OBSERVATION CHECKLIST

COURSE NAME/NUMBER_____DATE_____

TITLE OF TRAINING BLOCK_____ROOM NO

FILE # OF TRAINING BLOCK_____NO. OF STUDENTS

TIME SCHEDULED TO_____

TIME BEGAN_____TIME ENDED_____NO. OF BREAKS

INSTRUCTOR(S) NAME/DEPT_____

OBSERVER(S)_____

TEST WAS ADMINISTERED IN THE ROOM DESIGNATED IN THE SCHEDULE

* A COPY OF THIS CHECKLIST WILL BE FURNISHED TO THE INSTRUCTOR.

Comments:_____

_____

YES NO

1. Lesson Plan is available

a. Provides details of test administration procedures

b Is within the five year tolerance

2. Facilities Outside Inside

Classroom Classroom

NO YES NO

a. Classroom information displayed

Course Title

Class

(3) Section

(4 Subject

(5) Start/End Time

(6) Instructor(s) Name

3. Environmental Factors

a. Room Appearance/Maintenance _____

b. Noise Level _____

c. Student Control _____

d. Lighting _____

e. Temperature _____

f. Student/Equipment Ratio _____

g. Distractions _____

h. Adequate work space/seating per student

Comments: _____

_____

_____

4. Instructor

a. Student-to-Instructor Ratio_____

NO

b. Instructor taught subjects tested

c. Displayed a positive attitude.

d. Answered administrative data about the test

5. Administrative YES NO

a. Sufficient test materials were available.

b. Testing materials were distributed to student positions prior to students entering the area. ____ _____

c. Prior to the exam, students were provided with specific instructions. ____ _____

Administered by a qualified instructor

At least one instructor is present at all times.

6. Additional Comments:_____

_____

Instructor's Signature

APPENDIX E

GLOSSARY: ACRONYMS

AARTS	Army/American Council on Education Registry Transcript System
	American Council on Education
	Academic Evaluation Report
AF	United States Air Force
	Academy of Health Sciences, U.S. Army
AMEDDC&S	U.S. Army Medical Department Center and School

APFT	United States Army Physical Fitness Test
ARNG	United States Army Reserve National Guard
	Criterion Referenced Test
	Department of Academic Support
	Distinguished Honor Graduate
	Department of Training Development
	Defense System Network
	Enabling Learning Objective
EOCCT	End-of-Course Comprehensive Test
	Evaluation and Standardization Branch
	General Education Development
	National Guard Regulation
	Norm Referenced Test
	Systems Approach to Training
SEP	Student Evaluation Plan
	Standing Operating Procedure
	Terminal Learning Objective
TRAS	Training Requirements Analysis System
USAMEOS	United States Army Medical Equipment and Optical School
USAR	United States Army Reserve
USASAM	United States Army School of Aviation Medicine

MCCS-HE

FOR THE COMMANDER

OFFICIAL: CHARLES E DYER, II

Colonel, MS

Secretary of the General Staff

PAUL D. ANDERSON

CPT, MS

Adjutant General

DISTRIBUTION