Buros Standards for

Proprietary Testing Programs

Buros Standards for Proprietary Testing Programs

Preamble

The Buros Center for Testing at the University of Nebraska – Lincoln is an independent not-for-profit organization that has identified essential components for proprietary testing programs and established a mechanism to determine if these organizations meet these standards. The Buros Standards for Proprietary Testing Programs are intended for all legal entities that provide proprietary educational, psychological, credentialing, or personnel selection testing programs. These Standards are based on the *Standards for Educational and Psychological Testing*, jointly published in 1999 by the American Educational Research Association (AERA), the American Psychological Association (APA), and the National Council on Measurement in Education (NCME), the *Guidelines for Computer Based Testing* published in 2001 by the Association of Test Publishers (ATP), and the *Test Adaptation Guidelines* and the *International Guidelines on Computer-Based and Internet Delivered Testing* developed by the International Test Commission (2000, 2005). These published standards often represent broad suggestions – the Buros Standards were developed as a way to operationalize the expectations of these professional organizations and are therefore, more specific.

Buros strongly encourages organizations that develop and/or administer proprietary testing programs to recognize and endorse these Standards. The Buros Standards for Proprietary Testing Programs provide the framework for universal acceptance of individuals who have taken examinations from a certification program. Buros uses these Standards to review the quality of proprietary testing programs. The judgments made from such reviews and the defensibility of such reviews is limited by the availability of information at the time of the review. In addition, not all of the Standards will apply to all testing programs.

Buros supports the position that training at various levels is critical to performance and operational application of appropriate principles and practices in educational, psychological, credentialing, and personnel selection testing arenas. However, training criteria are not incorporated into the Standards. This decision is based on the:

- recognition that effective training incorporates a variety of instructional techniques, approaches and delivery systems;

- understanding that emergence of new delivery technologies, such as distance learning and computer based instruction (CBI), increases accessibility and availability of training;

- appreciation for difficulty in developing an "equivalency standard" to assess content, scope and effectiveness of diverse training methods; and

- recognition that an assessment tool is the common component in many educational, credentialing, and selection programs.

Any organization wishing to have their testing programs reviewed against the Buros Standards (either for an audit or an accreditation) should review the Buros *Stage 1 Application* available from the Buros website (http://buros.org/audits-and-accreditation) and contact a Buros representative about the audit (or accreditation) process.

TABLE OF CONTENTS

Buros Standards for
Proprietary Testing Programs

Section 1:
PURPOSE OF THE TESTING PROGRAM

1.1 The testing program shall have a defined purpose and a clear description of the population(s) for which the test is appropriate.

1.2 A testing program responsible for measuring examinees within a given content domain has a responsibility to the examinees and to the public to ensure that its tests are properly used. This includes the testing program providing documentation of anticipated unintended uses of the test scores so that they are explicit with how their test scores should be used/interpreted.

1.3 The defined purpose(s) of the testing program shall be supported by a documented validity framework. The validity framework shall provide for multiple sources of validity evidence in support of each intended use of exam scores. This evidence may be drawn from theory, test development procedures, and empirical research and shall be integrated to provide a coherent justification of the use(s) of the test scores for the each intended purpose.

1.4 The testing program shall include an ongoing program of validity research by which it continually gathers evidence to support the intended use of the test scores.

1.5 The test developer shall provide evidence that the results or outcomes of the validity research are used to make improvements to the testing program.

Section 2:
STRUCTURE AND RESOURCES OF THE TESTING PROGRAM

Structure of the Testing Program

2.1 The testing program shall be incorporated as a legal entity (applies to the parent organization if the testing program is a subsidiary of another organization).

2.2 The testing program shall demonstrate that the relationship between the testing program and related association, organization, or agency (e.g., licensing board) ensures the independence of the testing program and its related functions.

2.3   If a testing program provides both education and testing, the testing program shall administratively and financially separate any education and testing functions that are specific to any high stakes decisions related to the testing function.

2.4   A testing program shall not have any conflict of interest with an organization from which it is seeking accreditation.

Resources of the Testing Programs
2.5   The testing program staff shall posses the knowledge and skills necessary to conduct the testing program or have available and make use of non-staff consultants and professionals to sufficiently supplement staff knowledge and skills. This may be documented through staff and consultant vitae and qualifications.

2.6   Each testing program shall provide evidence that its staff (or staff and consultants/contractors) meet the following professional requirements:

a. capability to conduct or otherwise use a legally defensible and psychometrically valid content specification methodology (e.g., job/practice analysis, theoretical framework, specified academic content objectives);

b. capability to develop psychometrically sound examinations;

c. capability to develop and implement thorough procedures for security of the item bank, printed, taped, or computerized examinations, and examinee scores;

d. capability to manage data commensurate with the requirements for effective processing, reporting, securing, and archiving of examinee examination scores;

e. capability to understand and willingness to abide by the principles of fairness and due process;

Section 3:
EXAMINATION DEVELOPMENT

3.1   Examination development procedures shall be based on the relevant standards found in the most recent edition of *Standards for Educational and Psychological Testing*, developed jointly by the American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, the *Guidelines for Computer Based Testing* developed by the Association of Test Publishers, the *Test Adaptation Guidelines* and the *International Guidelines on Computer-Based and Internet Delivered Testing* developed by the International Test Commission, and all appropriate Federal requirements (for example, Americans with Disabilities Act, Equal Employment Opportunity Commission

standards, Uniform Guidelines). Examinations shall be revised as needed to be in compliance with changes in the *Standards for Educational and Psychological Testing* or in any of the Federal requirements.

Job Analysis or Content Specification

3.2   The test developer shall provide clear documentation of the method used to define the content included in the exam. Included in this documentation shall be any analyses or judgmental procedures used.

3.3   The knowledge, skills, abilities, and judgments assessed by the testing program shall be determined through a psychometrically sound job/practice analysis or other content specification methodology.

3.4   The panel of subject matter experts providing judgments for the job analysis or content specification should be a demographically and technically representative group and possess significant experience with the content being measured.

   a. For the job or practice analysis, the representative group shall include persons with experience in the various commercial aspects of the industry related to the content area and shall be sufficiently diverse as to avoid cultural bias and ensure fairness in content according to all Federal requirements.

   b. For tests that measure content other than job-related content, the representative group shall include individuals who have experience with teaching or experience with the content in the context in which the testing is to be undertaken (e.g., the content specifications for a 4th grade reading test should be developed by individuals who are knowledgeable about what constitutes acceptable reading content and practices in reading in the 4th grade).

3.5   The test *content framework* shall be developed using a psychometrically sound methodology which translates the results from the job analysis or content specification methodology into a comprehensive list of the knowledge, skills, abilities, and judgments (KSAJs) or content areas and, if appropriate, cognitive dimensions required for the construct(s) of interest.

3.6   Detailed test *specifications* (e.g., weighting of different types of exam content) shall be derived from the job analysis or other content specification method and reflect the content domain and purpose of the exam. The examination specifications, consisting of percentage weights or number of items devoted to each content area, shall be available to examinees and to the public.

3.7   The testing organization is required to systematically evaluate practices (in the job field) and content in the appropriate content domain(s) to assure that the job analysis findings (or content specification methodology from which exam content was determined) remain current in accordance with the intended usage of the exam scores and appropriate for the development of examinations. The maximum length

of use for any job analysis or content specification is five years from the date of completion.

Item Development

3.8   The testing organization shall provide clear documentation of all the procedures used in the development and review of exam items.

3.9   Exam items shall be written by qualified persons, who have knowledge of the content domain and have participated in training on item writing.

3.10   All items shall undergo a rigorous review process based on criteria assuring the accuracy of the item key and response format, the item classification according to the exam specifications, the item readability, and the appropriateness of the cognitive level of the item. The review process shall also include a procedure to detect potential item bias.

3.11   A testing organization shall have a procedure in place to assure that a sufficiently large item pool is maintained that appropriately represents the content included in the test framework.

Pilot Testing

3.12   The test developer shall provide clear documentation of the procedure(s) used to pilot test the exam items.

3.13   The test developer shall pilot-test any potential exam items to determine the appropriateness for inclusion in the operational form. Pilot tests shall yield data appropriate for determining the psychometric adequacy of an item to be included in the operational form.

3.14   The pilot sample shall be of adequate size and representative of the intended population of examinees with respect to ability (on the construct being measured) and relevant demographic characteristics (e.g., gender, ethnicity, geographic location).

3.15   The conditions under which the pilot data are gathered shall be the same as those by which the operational items will be administered.

3.16   If the exam is computer/internet based, the pilot test process should include a test of the system capabilities including capacity of the computer/internet server.

Creation of Final Exam Forms

3.17   The test developer shall provide clear documentation of the process used to assemble the final operational exam forms.

3.18 Item analyses shall be conducted with the pilot data to determine which items from the item pool should be included in the operational exam forms. Such analyses should include, at a minimum, evaluating item difficulty and item discrimination.

3.19 Items should be chosen for inclusion on the final forms based on adequate psychometric properties and fit to the test framework

3.20 The final exam forms should be evaluated for fit to the test framework and exam specifications by matching content to the framework and ensuring appropriate weighting of content on the exam.

Periodic Review

3.21 The testing program should provide documentation of the periodic test form review process including the results of this review and any actions that were taken based on the results.

3.22 Depending on the number of examinees and the number of possible administrations, at least annually, the testing organization shall conduct a psychometric review of its examination(s). The report should include the following summary statistics for all examinations administered since the last reporting period:

Test Characteristics (report by form)

a.   number of examinations administered;

b.   exam score descriptive statistics (mean, median, standard deviation, & range)

c.   reliability and (if appropriate) decision consistency coefficients;

d.   number and percentage of examinees passing the examination (if passing scores have been set), including separate reporting for original examinees and for retakes;

e.   number and percentage of examinees classified into performance classifications (other than pass/fail) by the examination (if passing scores have been set with the intent of making such classifications), including separate reporting for original examinees and for retakes

f.   model fit (if using IRT for data calibration).

Item Characteristics

g.   item discrimination (e.g., point-biserial correlation, discrimination indices, a-parameters)

      h.   item difficulty (e.g., p-values, b-parameters)

Test Adaptation

3.23 If a test is adapted from one language to another, the test developer shall provide clear documentation of the process by which the adapted version was created including documenting who was involved in the process, the process used to create the adapted version of the exam, and the results of the process.

3.24 When adapting an existing exam for use by a population of another language or culture, a test developer shall consider the impact of linguistic and cultural differences related to test techniques, test conventions, directions, item formats, item content, rubrics, and stimulus materials. To avoid potential problems in translation or adaptation of terms specific or idiomatic to the content domain, this process should be accomplished with the consultation of personnel competent in the languages of both the original and the translated version of the examination.

3.25 A test developer shall provide validity evidence to support the intended interpretation of test scores for the adapted version of the test specific to the population of interest. This includes evidence of content equivalency across languages.

Section 4:
EXAMINATION ADMINISTRATION

Application and Registration

4.1 The testing program shall provide clear documentation of the process by which exam candidates apply to take the exam. The documentation shall include application instructions given to candidates and procedures for review of applications.

4.2 The testing program shall have a process in place for review of applications by which they assure that all candidates who are permitted to sit for the exam meet the eligibility requirements including eligibility to receive special accommodations.

4.3 Eligibility requirements shall be documented and shall specify necessary qualifications and/or demographic characteristics definitive of the population of test-takers for whom the test is intended.

4.4 Examinations shall be scheduled far enough in advance to allow for timely shipment of supplies. Testing organizations may set specific time periods for required notification of a scheduled examination.

4.5   The testing program shall have in place a documented re-take policy specifying under what conditions and after what period of time an examinee may re-take the same exam.

Accommodations
4.6   The testing program shall provide clear documentation of the process by which they determine eligibility for a candidate to receive special accommodations that is compliant with the guidelines set forth by the Americans with Disabilities Act.

4.7   The testing program shall have clear instructions to candidates about the necessary documentation they are required to present to the testing program to receive special accommodations and the eligibility criteria for receiving accommodations.

4.8   Where special accommodations shall be made for eligible/otherwise qualified examinees under provisions of the Americans with Disabilities Act, arrangements shall be such that the security of the examination contents is not compromised. The examination contents shall not be revealed to any test administration personnel with any conflict of interest. A written affirmation to that effect and a written nondisclosure statement from the individual who was chosen to assist the otherwise qualified candidate shall be provided to the testing organization.

Administration Sites
4.9   The testing program shall offer the examination at a sufficient number of sites to ensure reasonable ease of access for at least a majority of candidates to the exam administration site.

4.10   If the testing program is administering a computer-based or internet-based exam, the testing program shall ensure the administration site conforms to all necessary hardware and software requirements.

4.11   Sites chosen for administering examinations shall conform to all legal requirements for safety, health, and accessibility for all qualified examinees. Additionally, the accommodations, lighting, space, comfort, and workspace for taking the examination shall allow all examinees to perform at their highest level of competency. Requirements at each site include but are not limited to:

   a.   accessibility in accordance with requirements of the Americans with Disabilities Act shall be available for all qualified examinees, whether it be the main site for an administration or an alternative site meeting all other requirements of the main site;

   b.   adherence to all fire safety and occupancy codes of the jurisdiction in which they are located;

c.   sufficient spacing between each examinee in the area in which the actual testing is conducted, or other appropriate and effective methods shall be in place to preclude any examinee from viewing another examinee's test.

d.   acoustics that allow each examinee to hear instructions clearly, using an electronic audio system if necessary;

e.   adequate lighting at each examinee's work space for reading fine print;

f.   ventilation and temperature appropriate for health and comfort of examinees;

Test Administrators and Proctors

4.12   Testing organizations shall provide documentation of the responsibilities of test administrators and of monitors/proctors and the minimum criteria for approval of test administrators and for monitors/proctors.

4.13   Responsibilities, duties, qualifications of test administrators and monitors/proctors shall be directed toward assuring standardized, secure examination administration and fair and equitable treatment of examinees.

4.14   Testing programs shall provide suitable programs of training for test administrators and monitors/proctors to enable them to maintain a standardized administration environment, and prepare them to deal with any issues, including security concerns that arise during administration, which could compromise the validity of test score interpretation.

4.15   The number of approved monitors/proctors assigned to a test administrator shall be sufficient to allow each examinee to be observed and supervised to assure conformance to security requirements. There shall be no less than one test administrator/proctor/monitor per thirty-five examinees for paper and pencil administrations, and per fifteen examinees for tests administered in electronic formats.

4.16   When an instructor/educator/trainer of a training or educational program administers, proctors or monitors an examination, the testing program shall have the proctor sign a confidentiality statement assuring they will not record or share any of the test information.

Procedures for Administration

4.17   The testing organization shall provide each test administrator with a manual detailing the procedures and requirements for all aspects of the examination administration process.

4.18   The testing organization shall present the examinee with detailed instructions specifying how the examinee will proceed through the exam, and clearly state any

rules or regulations a test examinee is expected to follow during the administration of the exam.

4.19 The test administration should be conducted in a standardized fashion according to the procedures documented in the test administration manual. Any disruptions to the procedure should be carefully documented by the test administrator/proctor and reported to the testing organization.

Section 5:
Scoring and Score Interpretation

Scoring

5.1 The testing program shall have clear documentation describing the procedure used for scoring an exam including a procedure to verify the accuracy of exam scores.

5.2 For paper and pencil tests or tests that use a separate answer sheet, completed answer sheets shall be shipped to the testing organization within a reasonable timeframe.

5.3 If the testing program requires the use of raters or subjective scorers for any part of the exam, the testing program should provide adequate training for the scorers.

5.4 If the testing program requires the use of raters or subjective scorers for any part of the exam, the testing program shall have in place a process by which the accuracy of these scores are checked on an ongoing basis (e.g., double scoring system).

5.5 If the testing program uses a computer-based scoring program to compute exam scores, the testing program shall have in place a process to verify the accuracy of these scores.

Scaling

5.6 Clear documentation shall be provided for all procedures related to scaling of the exam scores including development of the scale, procedures used to translate raw scores to the reporting scale, and rationale for scaling method.

5.7 The development of the score scale and the translation of raw scores to scale scores shall be based on psychometrically sound procedures appropriate for the intended interpretation of test scores.

5.8 Examinees should be provided with adequate documentation to interpret scale scores for norm-referenced or criterion-referenced tests.

Standard Setting

5.9   The procedure and results of the standard setting activity should be clearly documented including the method by which recommended passing standards were determined, the recommended passing score standard(s), an estimate of the panelists' variability in their recommendations for the passing standards, and the panelists' perceptions of the process.

5.10   Performance standards shall be set using an appropriate standard setting method that has been field tested with similar exams.

5.11   Standard setting panelists shall be subject matter experts who are familiar with the target population of examinees and the construct/content being assessed. The group of panelists selected to participate in the standard setting should be representative of the stakeholders who are qualified to make decisions about the required proficiency of examinees.

5.12   Panelists shall be provided a clear description of the purpose of the standard setting workshop including the intended use of the cut score being estimated and definition of the performance standard(s). Panelists shall be provided adequate time to discuss the purpose of the cut score and the performance standard(s) to ensure a common understanding of these concepts among panelists.

5.13   The standard setting process should include adequate training by which panelists can become familiar with the process they will use to set a passing standard recommendation. This training should include practice with the standard setting method.

5.14   The standard setting process should include an opportunity for panelists to provide feedback on the standard setting process and their comfort and confidence in the estimated cut scores.

Norm Referenced Interpretation

5.15   The documentation of the norming study shall provide a clear description of the process including a description of the sample, administration format, test dates, and results including sample weightings and descriptive statistics of the norms.

5.16   If the testing organization utilizes norm-referenced data for reporting examinee scores, a norming study shall be conducted to collect appropriate data on which to base score interpretation.

5.17   The selection of individuals from the norming population to be included in the norming study shall be based on psychometrically sound sampling procedures. The sample selected shall be representative of the target population with respect to demographic and other relevant characteristics.

5.18   The norming study should be conducted in the same conditions in which the operational exam will be administered.

Equating/Comparability

5.19   The process by which equivalence across forms, languages, etc. is determined for multiple test forms shall be clearly documented as public information.

5.20   When the examination is administered in more than one format, evidence shall be collected and provided to assure that all knowledge and competencies are assessed in a manner that will result in reliable and valid interpretation of examination scores across formats. Evidence of equivalence with other examination formats shall be provided.

Score Reporting and Record Keeping

5.21   The testing organization shall prepare score reports that include a guide to interpreting any scores reported (e.g., sub-scores in addition to total scores). Score interpretations presented in the guide shall be in accordance with the test's defined purpose and intended uses of the scores.

5.22   The testing organization shall have a procedure in place by which score reports will be delivered to authorized recipients in a time frame specified in the application for accreditation. If there is a delay due to problems in verification or authentication of scores, examinees or appropriate authorities should be informed and an approximate date for release of the scores should be announced. The testing organization should have ongoing communication with examinees, or appropriate authorities, and with the test administrator until the scores are verified and released.

5.23   The testing organization shall provide evidence that examinees failing the examination are given information on general areas of deficiency if these general areas have sufficient evidence of reliability for the intended inference.

5.24   The testing organization shall maintain record of test administrations and score reports for a period of no less than 5 years.

5.25   If a credential, certificate, or similar recognition is issued as a result of successfully meeting a predefined performance level on an examination, records, including an electronic listing of examinees, maintained by testing programs shall identify the examination form and version (if appropriate) and specify the date the examination was taken.

5.26   The testing organization shall provide documentation of any procedures ensuring the confidential release of examinee scores to the examinee and other authorized persons or organizations as indicated in the policy documents of the testing program (e.g., school, licensing agency).

5.27 If a testing program is considered to be high stakes for examinees, the testing organization shall have in place a formal appeals procedure by which examinees can question eligibility to take the exam or review their examination scores. This process shall include a procedure by which exam scores will be investigated and scoring issues addressed.

Section 6:
EXAM SECURITY

6.1 The testing organization shall have procedures in place to assure the security of all test materials (e.g., items, forms, etc). This includes any steps taken during the development of the examinations.

6.2 The examinations shall be administered in a manner that maximizes the security of the exam contents. This includes administering examinations in a manner that allows absolutely no one other than the examinees to see the contents of the booklet or alternative medium, both before and after the examination is administered

6.3 When tests are administered in a paper and pencil format, security of the examination materials shall be maintained in shipments to and from the examination administration site.

6.4 For exams administered via the Internet, test publishers shall provide for appropriate security measures that protect the security of the test materials as well as the security of any candidate responses and test scores.

6.5 The testing organization shall have security procedures in place to protect the security of the exam items, forms, and administration records wherever such materials and information are stored.

6.6 The testing organization shall provide procedures to be followed in any instance where the security of an examination is, or is suspected, to be breached. Included shall be specific procedures for handling, and reporting to the accrediting organization, suspected or alleged cheating incidents, lost or stolen booklets, intentional or unintentional divulging of test items by examinees or test administration personnel, or any other incidents perceived to have damaged the security of the examination or any of its individual items.

Section 7:
RESPONSIBILITIES TO EXAMINEES AND THE PUBLIC

This section identifies specific responsibilities a testing program has to examinees and to the public. Many of these requirements overlap with other Standards and where appropriate, these connections are identified.

7.1 Responsibilities to Examinees. A testing organization shall:

a. Not discriminate among examinees as to age, sex, race, religion, ethnic origin, disabilities or marital status and shall include a statement of non-discrimination in announcement of the testing program.

b. Provide all examinees and appropriate authorities, with copies of formalized procedures for application for the testing program (Standard 4.1).

c. Have a formal policy for the periodic review of application and examination procedures to ensure that they are fair and equitable (Standard 4.12).

d. Provide evidence that competently proctored testing sites are readily accessible (Standards 4.9, 4.11, 4.14, 4.15).

e. Provide evidence of uniformly prompt reporting of examination results to examinees or appropriate authorities (Standard 5.22).

f. Provide evidence that examinees failing the examination are given information on general areas of deficiency if these general areas have sufficient evidence of reliability for the intended inference (Standard 5.23).

g. Provide evidence that each examinee's examination results are held at the appropriate level of confidentiality (Standards 5.26, 6.3 and 6.5).

h. For high stakes programs, have a formal policy on appeals procedures for examinees questioning eligibility or any part of the testing program (Standard 5.27).

7.2 To take an examination, an individual shall meet the eligibility requirements defined by the testing program (Standard 4.3).

7.3 If a credential, certificate, or similar recognition is issued as a result of successfully meeting a predefined performance level on an examination, records, including an electronic listing of examinees shall be maintained by the testing program that identify the examination form administered and the date the examination was taken (Standard 5.25).

7.4   For testing programs used in credentialing, a testing organization shall maintain a registry of individuals who have taken their examinations. Any credential, certificate, or similar recognition awarded by the testing organization shall not be confusing to employers, consumers, related professionals and/or other interested parties.

7.5   Testing programs shall have a published protocol for systematically investigating problems presented by users of the Program, including specific concerns about examination items, administration procedures, treatment of candidates, or other matters involving potential legal defensibility of the examination or program. The protocol will include a published time frame for reporting findings to the User.

DEFINITIONS

Accreditation means that an independent accrediting organization has reviewed a proprietary testing program and has verified that it meets standards set by the Buros Center for Testing.

Accredited examination means Buros acknowledgement that an examination has been reviewed and determined to have met the test development and administration portions of the Standards for Accreditation of Proprietary Testing Programs.

Accrediting organization The accrediting organization is the Buros Center for Testing.

Accredited program means a proprietary testing program that has been evaluated and listed by Buros and has met the Buros Standards for such programs.

    (a) refers to a designation based upon an independent evaluation of factors such as the testing program's purpose; organizational structure; organizational resources; policies; public information regarding program scope, maintenance requirements, discipline and appeals procedures; and test development, psychometric evidence, administration, and security. The specific factors depend on the nature of the testing program.

    (b) does not refer to training functions or educational programs.

Competency means a defined combination of knowledge, skills and abilities required in the satisfactory performance of a job or the attainment of an educational outcome or standard.

Competency examination means an instrument that assesses whether an individual has attained at least a predefined level of competency. It shall be based on a thorough analysis of requirements and specification of necessary content.

Demographic data means the statistical data of a population, especially the data concerning age, sex, ethnic distribution, geographic distribution, education, or other information that will describe the characteristics of the referenced group.

Educator, in this instance, means an instructor in a program or a course of study that includes content specific competencies.

Elements of Knowledge mean those specific skills and knowledge necessary for an examinee to perform in a minimally competent manner.

Entry level performance means the minimal knowledge and skills needed to perform effectively for a license, specialization area (in the case of certification), job, or

academic activity (e.g., grade level or admission to an educational program), but not necessarily beyond that level.

Equivalency means that there is psychometric evidence that various forms or versions of an examination cover the same content and their respective scores (especially passing scores, when appropriate) represent the same degree of competence.

Examination forms mean alternate sets of test questions to assess the same competencies, conforming to the same examination specifications.

Examination specifications mean the description of the specific content areas of an examination, stipulating the number or proportion of items for each area of competency and the level of complexity of those items. The specifications may be based on a job analysis or other accepted methodology for content specification. In educational settings the examination specifications may be a set of content area objectives or standards. Psychological test specifications may be based on a theoretical description of a construct, or it may be derived empirically based on clinical practice or experience (or by such documentation as the more current edition of the Diagnostics and Statistical Manual of the American Psychological Association).

Examination version means a test in which the exact set of items in an examination form is presented in another order, language, manner, or medium.

Item bank means all of the items that have been developed for multiple forms of an examination. It includes all of the items available to create examination forms.

Job Analysis/Content Specification means the description of functions or tasks required for an individual to perform to entry level standards in a specific job, profession, or set of competencies, including information about the attributes required for that performance. It defines the performance dimension of a job, profession, or set of competencies, and includes knowledge, skills and abilities necessary to carry out the tasks or functions associated with the purposes of the test (e.g., college admissions, grade level promotion, determination of achievement levels).

    (a) tasks are the individual functions, whether mental or physical, necessary to carry out an aspect of a specific job, profession, or set of competencies

    (b) knowledge, skills and abilities (KSA's) include the information and other attributes that the examinee shall possess. They include information and understanding as well as learned behaviors and natural attributes.

Legal entity means an organization structured in a manner that allows it to function legally and be recognized as a responsible party within the legal system

Legally defensible means the ability to withstand a legal challenge to the appropriateness of the examination for the purpose for which it is used. The challenge may be made by actual or potential examinees or on behalf of the public. Examinee's challenges may pertain to perceived bias of the examination or inappropriately chosen content. Challenges on behalf of the public may claim that the examination does not provide adequate measures of an examinee's knowledge, skills and abilities within a given content domain.

Monitor (See Proctor) means an authorized agent of the program sponsor who ensures the security of the examination and verifies the identity of examinees.

Proctor means persons under the supervision of a test administrator, assisting by assuring that all aspects of an examination administration are being carried out with precision, with full attention to security and to the fair treatment of examinees. Proctors have the responsibility and shall have the ability to observe examinee behaviors, accurately distribute and collect test materials, and assist the test administrator as assigned. They shall have training or documented successful experience in monitoring procedures and shall affirm in writing an agreement to maintain test security and to assure that they have no conflict of interest.

Program sponsor means an entity that requires the test be administered for the specified purposes of the test. A program sponsor may be an employing organization, a certifying organization, a state government licensing agency, an educational organization (e.g., school district, institution of higher education), or a program that promulgates a test that may be used voluntarily to achieve the purposes of a program sponsor (e.g., an organization conducting a training course related to the testing program).

Psychometric means scientific measurement or quantification of human qualities**,** traits or behaviors.

Psychometrician means a professional with specific education and training in development and analysis of tests and other assessment techniques and in statistical methods. Qualifications may vary but usually include at least a bachelor's degree and a minimum of two formal courses in tests and measurement (including test development) and a minimum of two courses in statistical methods.

Reliability means the degree to which scores from a measure are free from random error. Reliability may be estimated using appropriate methods (e.g., test-retest, alternate form, internal consistency, estimates of inter-rater agreement, decision consistency reliability, or other measures of generalizability of scores).

Test Administrator means the individual at the test site who has the responsibility for conducting an examination. Test Administrators shall have training, documented successful experience, or a combination of experience and training in test

administration and security procedures. They shall provide written assurance of maintaining confidentiality of test contents and of adherence to standards and ethics of secure examination administration.

Trainer/educator, in this instance, means a professional with appropriate expertise who conducts a course in a content area for examinees in that content area.

Validity means the extent to which a score from a test or other type of assessment permits correct inferences regarding examinees in a given content area. Specifically, does the test produce scores that possess sufficient supporting evidence for their intended use(s)?