

APPENDICES

The following terms and concepts denote important aspects of employment testing. Definitions represent a common understanding among professionals in the assessment field and are based upon professionally recognized principles and standards.

- ability** A present competence to perform an observable behavior or a behavior which results in an observable product.
- adverse impact** A substantially lower rate of selection in hiring, promotion, or other employment decisions that works to the disadvantage of members of particular race, ethnic, religious, gender, or age groups, which are recognized as protected against illegally discriminatory employment practices.
- assessment** The measurement of an individual's employment or job-related qualifications or characteristics using a variety of testing procedures, including, for example, written examinations, performance-based tests, interviews, evaluations of training and experience, reference checks, and medical evaluations.
- assessment center** A method of evaluating candidates through the use of a battery of test procedures which assesses the level of expertise and possession of requisite knowledge, skills, and abilities. Test components comprising an assessment center may include structured interviews, role-play exercises, writing exercises, group exercises, and written examinations. Assessment centers may be used for a multitude of job classifications; however, they traditionally have been used for managerial/supervisory assessment.
- certification** The process by which candidates on an eligible list are referred to hiring supervisors for consideration in filling vacant positions in the job classification for which the eligible list has been created.

chance score	The score that a candidate could obtain on a test by purely guessing.
cognitive test	A measure of an individual's general mental ability or intelligence.
compensatory scoring model	A scoring model which assumes that high scores on one or more predictors (i.e., tests) can substitute or compensate for low scores (or scores below the threshold of minimal acceptable competence) on other predictors.
construct	A theoretical characteristic or concept (e.g., numerical ability, conscientiousness) that has been constructed to explain observable patterns of behavior.
construct validity	The extent to which a test measures a specific theoretical construct, characteristic, or trait identified as important for job success.
content validity	The extent to which the content of a test (1) samples or represents the subject area or behavior it is intended to measure and (2) is related to requirements and qualifications important for job success.
converted score	A raw score that has been converted by numerical transformation (for example, to a standard score) to facilitate comparison of individual scores with group norms.
correlation	A statistic that indicates the degree to which two variables relate to each other, such as a test score and job performance, or one test with another test.

criterion	A measure of performance, such as productivity rate, accident rate, or supervisory ratings. Test scores (<i>also known as predictors</i>) are used to predict criteria.
criterion-referenced assessment	The assessment of candidate qualifications to determine competence in a specific content domain (e.g., math skills, supervisory principles, technical terminology); scores derived from criterion-referenced tests indicate levels of mastery as determined by a standard rather than by comparative data across candidates.
criterion-related validity	The degree to which scores on an assessment instrument correlate with some external criterion, such as job performance. Criterion-related validity is an empirical or statistical relationship between test performance (i.e., scores on a predictor) and job performance (i.e., scores on a criterion measure). When the assessment instrument and the criterion are measured at about the same time, it is called concurrent validity; when the criterion is measured at some future time, it is called predictive validity.
distribution	A tabulation of scores from high to low, or low to high, showing the number of individuals obtaining each score or falling in each score interval. (<i>Also known as a frequency distribution.</i>)
eligible list	A list documenting the results of a civil service examination process; the eligible list is comprised of those candidates who successfully complete the examination process and are eligible for appointment in the job classification for which the examination process was conducted.
equal employment opportunity	Providing equal access to jobs, work assignments, training, and other employment-related opportunities to all job applicants, candidates, and employees, regardless of political affiliation, race, ethnicity, color, ancestry, national origin, religion, gender, sexual orientation, age, marital status, medical condition, or mental or physical disability.

examination process	A formal testing process conducted to assess candidate job-related qualifications, which results in the establishment of an eligible list from which appointments to positions in the civil service may be made; an examination process may be comprised of any number of selection procedures deemed appropriate for the assessment of the requisite job-related qualifications, including, for example, written examinations, interviews, performance tests, work sample exercises, and supplemental applications.
high-fidelity test	A test which simulates actual tasks of the job and requires candidates to respond by behaving as if they were on the job.
in-basket exercise	A work sample test in which candidates are provided with a variety of materials of varying importance and priority typically handled on the job. Candidates are given a pre-determined period of time in which to review, prioritize, and make decisions about handling the materials. Candidates may respond to the in-basket items by actually taking action or indicating what actions they would take if on the job. Candidate performance is scored using pre-determined scoring criteria; the scoring process may be accomplished either through a paper review of the candidates' actual or indicated actions or an interview with the candidates requiring that they explain the actions they would take.
item analysis	An objective, empirical, statistical process yielding quantitative information about the effectiveness of a test and the individual items which comprise the test. The item analysis provides information regarding the difficulty of each test item, the relative attractiveness of each of the item alternatives, and the extent to which each item discriminated among the candidates with respect to a chosen criterion.
job analysis	A systematic method of studying jobs to identify the tasks, duties, responsibilities, and working conditions associated with the jobs and the knowledge, skills, abilities, and other characteristics required for successful job performance.

job-relatedness	A standard met when the knowledge, skills, abilities, and characteristics assessed in a selection procedure are demonstrated through an analysis of the job to be necessary for successful job performance.
KSA	An acronym meaning knowledge, skill, or ability.
knowledge	A body of information applied directly to the performance of a function.
linear transformation	A mathematical process whereby raw scores are rescaled (or recalibrated) to a different numerical range; all properties of the original distribution of raw scores are retained in the distribution of the new rescaled or recalibrated scores.
low-fidelity test	A test which simulates a job task in a written or verbal description and elicits a written or verbal response from candidates rather than an actual behavioral response.
low-fidelity written simulation	A written examination in which candidates are provided with hypothetical work situations and asked to indicate how they would deal with such situations or what actions they would take in response to the hypothetical situations. (<i>Also known as a situational judgment test.</i>)
mean	The average score in a group of scores, computed by adding all the scores and dividing the sum by the number of cases.
median	The middle score in a group of ranked scores. It is the point or score that divides the group into two equal parts. The median is also known as the 50 th percentile.

mode	The most frequent score in a group of ranked scores.
motor work sample test	A work sample test which involves the physical manipulation of things.
multiple cut-off scoring model	A noncompensatory scoring model comprised of multiple test parts which requires a candidate to pass all of the test parts or components in order to be successful in the overall testing process; a model used when some minimal level of proficiency on one or more tests is required for job success.
multiple hurdle scoring model	A noncompensatory scoring model comprised of multiple test parts which requires a candidate to pass each of the test parts or components in sequence in order to advance to the next test component and ultimately be successful in the overall testing process; a model used when it is advantageous to select only those candidates who meet minimal levels of proficiency to continue in the selection process, such as when subsequent test components may be complex, lengthy, or expensive to administer. <i>(Also known as successive hurdles model and sequential-decision strategy.)</i>
narrative supplemental application	A supplemental application in which candidates provide narrative responses to a series of job-related questions regarding their previous work experience, education, training, and job-related qualifications. A panel of raters evaluates the candidates' responses to each of the questions using pre-defined scoring criteria. <i>(Also known as the behavioral consistency method.)</i>
non-cognitive test	A measure of an individual's non-cognitive traits and characteristics, such as personality, values, attitudes; typically, non-cognitive tests are used in conjunction with other selection procedures to assess a candidate's job-related qualifications, not as the sole selection procedure.

noncompensatory scoring model	A scoring model which assumes that low scores (or scores below the threshold of minimal acceptable competence) on one or more predictors (i.e., tests) can not be substituted or compensated for by high scores on other predictors.
norm-referenced assessment	The assessment of candidate qualifications to determine relative comparisons among all candidates tested; scores derived from norm-referenced tests are interpreted by comparing them to the scores obtained by the other candidates taking the same test rather than interpreting based upon a criterion-referenced standard of performance.
normal curve	A mathematical curve that is the basis of many statistical analyses. The curve is bilaterally symmetrical, with a single bell-shaped peak in the center. Most distributions of human traits, such as height, mathematical ability, and manual dexterity, approximate the normal curve.
norms	Descriptive statistics that are used to summarize the test performance of a specified group, such as a sample of workers in a specific occupation. Norms are often assumed to represent a larger population, such as all workers in an occupation.
objectively scored written examination	A written examination in which candidates are asked to respond to a variety of items to which there is only one correct answer. Objectively scored item types include multiple-choice items, true-false items, and matching items. Examination answer sheets are generally machine scored.
oral presentation	A work sample test in which candidates are asked to prepare verbal presentations, typically of a specific length on a specifically indicated topic. It is common for candidates to be provided with job-related materials and documents to review, analyze, and upon which to base their presentations. The presentations are observed and evaluated by raters using pre-determined scoring criteria.

pass point	The minimum score that a candidate must achieve in order to be successful in a testing procedure; the pass point on a test typically equates to the minimal acceptable level of competence required on the job as measured by performance on the test. <i>(Also known as a cutoff score.)</i>
percentile score	The score on a test below which a given percentage of scores fall. For example, a score at the 80 th percentile is equal to or higher than the scores obtained by 80% of the people who took the test.
performance test	A category of motor work sample tests where there is some physical activity or the physical manipulation of things.
personal characteristics	An identifiable aspect of a person's physical makeup or typical behavior pattern that contributes to successful job classification performance.
portfolio	A testing procedure in which candidates provide and discuss job-related examples of their visual, written, oral, or artistic achievements. Scoring is based upon pre-determined scoring criteria.
predictor	A measurable characteristic used to predict criterion performance, such as test scores or interview panel ratings.
pre-exposed interview questions	Candidates are given a pre-determined amount of time immediately prior to their interview to review and take notes on one or more of the structured interview questions. Candidates then respond to the questions during their interview; in some instances, candidates may be allowed to reference their notes when responding to the pre-exposed questions.

pre-interview exercises	The use of one or more performance-based exercises in the structured interview process. In addition to being asked structured interview questions, candidates are asked to perform a task or series of tasks or produce a product. The intent of a pre-interview exercise is to combine a performance test with a structured interview.
psychomotor	Pertaining to bodily movements associated with mental processes.
range	The difference between the highest and lowest scores obtained by a group on a test or selection procedure.
rank ordering	The process of ranking individuals based on their relative test scores, from the highest to the lowest score.
raw score	The obtained score on a test, usually determined by counting the number of correct answers or totaling the number of points scored.
reliability	The degree to which test scores are dependable, repeatable, and unaffected by measurement error; reliability refers to the consistency of measurement.
reliability coefficient	A coefficient of correlation that indicates the degree to which test scores are consistent; that is, dependable, repeatable, and unaffected by measurement error.
role-play exercise	A work sample test in which candidates assume the role of incumbents in the job. Candidates interact with trained role players in hypothetical, job-related situations. The interaction is observed and evaluated by raters using pre-determined scoring criteria.

scoring model	The fundamental design of an examination process which includes (1) the identification of the selection procedures to comprise the examination process and (2) the determination of how each selection procedure will be scored and weighted.
selection procedure	Any test or procedure used to measure an individual's job-related qualifications or characteristics. Selection procedures include the full range of assessment techniques from written examinations, performance-based tests, interviews, evaluations of training and experience, reference checks, medical evaluations, to probationary periods.
selection process	The process by which applicant/candidate qualifications are measured against job-related requirements to identify those applicants/candidates who possess the requisite qualifications to perform successfully in a given job. The State's civil service selection process includes recruitment and advertising of employment opportunities, testing for eligible list establishment, effecting hiring decisions based upon hiring interview processes and any additional testing of candidate qualifications that might be required, and administering the required civil service probationary period.
selection rate	The proportion of applicants or candidates who are hired, promoted, or otherwise selected based upon the results of a selection process or individual selection procedure.
self-rating supplemental application	A supplemental application in which candidates rate their previous work experience, education, training, and job-related qualifications using point scales or scaled responses to indicate their level of expertise or possession of requisite knowledge, skills, or abilities. The candidates' responses may be machine scored or hand scored. <i>(Also known as the point method.)</i>

skill	A present, observable competence to perform a learned psychomotor act.
standard deviation	A statistic used to describe the variability within a set of scores. It indicates the extent to which scores vary around the mean or average score.
standard error of measurement (SEM)	A statistic that gives an indication of the amount of error in a measurement system. It indicates a range within which a test taker's "true" score is likely to fall.
standard score	A score that describes a candidate's raw score in terms of its distance from the mean score in standard deviation units.
standardized test	A test developed using professionally prescribed methods and providing for specific administration requirements, instructions for scoring, and instructions for interpreting scores.
structured interview	A formal, systematic, and standardized interview process in which all candidates are asked the same pre-established, job-related questions in the same order. The interview panel evaluates the candidates' responses using a pre-defined scoring guide consisting of benchmark answers and anchored rating scales.
subject matter expert (SME)	A person who is considered to be a job-content expert and participates in one or more phases of the selection process. SMEs are typically incumbents in the job classification for which the selection process is conducted and first-level supervisors.

subjectively scored written examination	A written examination in which candidates are asked to respond to questions in an open-ended fashion. Subjectively scored item types include essay, short answer, and fill-in-the-blank questions. Candidate responses to subjectively scored written examination questions are hand-scored using pre-determined scoring criteria.
supplemental application	An unassembled examination used to evaluate a candidate's job-related training, education, experience, accomplishments, and knowledge, skills, and abilities.
task	An activity performed in a job.
test	A systematic instrument or procedure that measures sample behavior or performance. The instrument should be systematic in three areas: content, administration, and scoring. A personnel or employment test is the general term for an assessment instrument used to measure an individual's employment qualifications, capabilities, or characteristics.
true score	A score entirely free of errors of measurement.
utility analysis	The bottom-line benefit to an organization for using a particular selection procedure. Utility analysis is a cost-benefit technique for selection programs.
validation	The investigative process of determining and documenting (1) <i>what</i> a test or selection procedure measures and (2) <i>the relationship</i> between scores from the procedure and some criterion measure, such as job performance.

validity	The degree to which a test or selection procedure accurately measures a body of knowledge, skills, abilities, and/or other characteristics identified as being required for successful job performance.
validity coefficient	A numerical index that shows the strength of the relationship between a test score and a criterion, such as job performance; derived from determining the statistical or empirical relationship between test performance and a criterion measure (e.g., job performance).
verbal work sample test	A work sample test which involves a problem situation requiring language skills (written or verbal) and may include interaction with people.
video-based test	A work sample test in which vignettes of on-the-job situations are depicted in video format on a screen, or computer monitor. Candidates are asked a series of questions about how they would react to or handle the situation vignettes. Candidates could be questioned via multiple choice items (which could be machine scored), open-ended essay or short answer questions, or raters in an interview situation. Scoring is based on pre-determined scoring criteria.
visual resume	A testing procedure in which candidates present and discuss relevant aspects of their experience, education, and training as they relate to the requirements of the job. Raters observe and evaluate the formal presentation using pre-determined scoring criteria.
work behavior	An activity performed to achieve the objectives of the job. Work behaviors involve observable (physical) components and unobservable (mental) components. A work behavior consists of the performance of one or more tasks.

work sample test	A broad class of assessment techniques that replicate job tasks or job behavior, are administered in a standardized manner, and allow for the observation of candidate behavior in performing the prescribed tasks.
writing skills assessment	A work sample test in which candidates are typically tasked with providing a written response to a specific situation or set of facts. Candidates may be provided with job-related materials and documents to review, analyze, and upon which to base their written response. Candidate responses may be scored on content, writing skill mechanics, or a combination of content and writing mechanics, and are hand-scored by subject matter experts using pre-determined scoring criteria.
written examination	A testing procedure in which candidates' job-related knowledge and skills are assessed through the use of a variety of item formats. Written examinations are either objectively scored or subjectively scored.