

SIOP Employment Testing Guide

Adapted from www.SIOP.org

The following guide to Employment Testing comes from the Society for Industrial and Organizational Psychology's website, SIO.org. This guide will answer questions that you have about using assessments for selection, competitive assessment; and the last section, which points end users to websites to find assessments, could be used to help you market your business.

SIO's Mission Statement

The Society for Industrial and Organizational Psychology is a Division within APA that is also an organizational affiliate of APS. The Society's mission is to enhance human well-being and performance in organizational and work settings by promoting the science, practice, and teaching of industrial-organizational psychology. Towards this end, SIO:

- Supports SIO members in their efforts to study, apply, and teach the principles, findings, and methods of industrial-organizational psychology
- Provides forums for industrial-organizational psychologists to exchange research, insights, and information related to the science, practice, and teaching of industrial-organizational psychology
- Identifies opportunities for expanding and developing the science and practice of industrial-organizational psychology
- Monitors and addresses challenges to the understanding and practice of industrial-organizational psychology in organizational and work settings
- Promotes the education of current and future industrial-organizational psychologists
- Promotes public awareness of the field of industrial-organizational psychology

Employment Testing

- 1. Overview.....4
 - What is an employment test?
 - When do employment tests make the most sense?
 - Reasons for testing
 - Reasons for NOT testing
- 2. How Many U.S. Companies Use Employment Tests?.....7
- 3. Establishing an Effective Employee Testing Program.....8
- 4. Types of Employment Tests.....11
 - Introduction
 - Assessment Centers
 - Biographical Data
 - Cognitive Ability Tests
 - Integrity Tests
 - Interviews
 - Job Knowledge Tests
 - Personality Tests
 - Physical Ability Tests
 - Work Samples and Simulations
- 5. Types of Test Formats.....18
 - Introduction
 - Pros and Cons of Multiple Choice Items
 - Pros and Cons of True/False Items
 - Pros and Cons of Essay Items
 - Guidelines for Using Multiple Choice or True-False Items
 - Guidelines for Using Essay Items
- 6. Effective Interviews.....21
- 7. Work Samples and Simulations.....24
- 8. Individual Psychological Assessment.....26
- 9. Information to Consider When Creating or Purchasing an Employment Test.....28
- 10. Seven Questions to Ask a Vendor Before Purchasing a Test.....31
- 11. Using Tests Effectively in the Public Sector.....34
- 12. Employment Testing Websites.....36

Employment Testing Overview

What is an employment test?

A test can be defined two different ways. From an assessment standpoint, a test is a standardized series of problems or questions that assess a person's knowledge, skills, abilities, or other characteristics. From a legal standpoint in the U.S., the *Uniform Guidelines on Employee Selection Procedures* (1978) defines a test as any method used to make an employment decision. Thus, a tool could be a test under one definition but not under the other (resume reviews are a test from a U.S. legal standpoint, but are typically not standardized in such a way as to be considered a test by most people involved in assessment). In this document, when we refer to employment testing, we mean standardized tools used in hiring, promotion, demotion, membership, referral, retention, and licensing and certification decisions.

There are different kinds of tests. Tests vary according to their mode of administration (e.g., paper and pencil vs. Web-based), their content (e.g., interpersonal skills, mathematical ability), their level of standardization or structure, their costs, their administrative ease, and many other factors. These factors are discussed in detail elsewhere (**for more information, see *Types of Employment Tests***).

When do employment tests make the most sense?

Employees affect an organization's performance and profitability. Hiring or promoting people who are unsuitable costs time, money, and potential new business. Carefully developed and administered employment tests can provide organizations with a way to decide systematically and accurately which people have the ability to perform well on the job, will not turnover, won't engage in counterproductive behaviors, or will be able to learn from training programs. Tests can also benefit individuals who are better matched to positions for which they are suited and in which they will wish to remain.

Reasons for testing

Some of the more commonly cited reasons for testing are:

Testing leads to savings in the decision-making process. Employment tests can be a cost effective way to pare down the applicant pool. Tests can make the decision process more efficient because less time is spent with individuals whose characteristics, skills, and abilities do not match what is needed. However, some tests do require more time up-front with individuals to determine who is and who isn't qualified. In these cases, tests can still result in savings from not training and compensating individuals whose productivity would be low or who would not remain on the job.

The costs of making a wrong decision are high. For certain employment decisions, a wrong decision can be very costly in terms of training costs, errors made by a poor performer, costs of replacement, etc. For these types of decisions, investing in testing may be seen as a particularly worthwhile endeavor if testing reduces the number of wrong decisions.

The job requires attributes that are hard to develop or change. Tests are often used for assessing characteristics that cannot be developed through training but are acquired over long periods of time or even a lifetime (e.g., personality traits, in-depth knowledge of a profession).

Hard-to-get information can be obtained more easily and efficiently. One important advantage of using employment tests is that they can often provide information about an individual that is not easily obtained using other methods, or that would be much more costly to obtain by other means.

Individuals are treated consistently. Using standardized tools in employment decision-making ensures that the same information is gathered on each individual and used in a similar way in decisions. Employers often turn to testing because of the unfairness of less standardized processes, in which individuals are not all treated in a similar way and similar information is not gathered on all individuals. Subjective biases can easily creep into decisions if the process for making decisions is unstandardized.

There are a lot of applicants. Sometimes the sheer number of individuals to consider for an employment decision leads an employer to choose testing as the most efficient and fair means of making a decision in a timely manner.

Reasons for NOT testing

Some of the most commonly cited reasons for not testing are:

Costs. While tests vary in their costs (e.g., developing customized tools costs more than purchasing off-the-shelf products, extensive assessments typically cost more), the cost of testing may be easily offset when considering costs of low productivity, errors, retraining times, and turnover. For example, conservative estimates of the cost of turnover range from 1/3-1/2 of the annual salary of the employee that is being replaced. The costs of replacing management, executive and highly skilled talent can easily be 1-2 times the annual incumbent's salary. Further, the costs associated with hiring a "wrong" employee who makes mistakes can be quite high. Testing can be a valuable investment for organizations to make in hiring and retaining talent.

Fear of legal action. Sometimes concerns are raised about the legality of using tests in hiring. As with any other method of making employment decisions, tests can be scrutinized if there is a belief that discrimination in employment decisions has occurred. Adverse impact exists when the selection rate of a given demographic group (e.g., females vs. males, whites vs. blacks, etc.) is substantially lower than the selection rate of the majority group. While any selection procedure may show score differences that result in exclusionary effects upon a group, some types of tests (e.g., physical ability, cognitive ability) are more likely to show such score differences. Despite these differences, these tests are often accurate predictors of job performance and other outcomes of interest. Before using a test, it is important to anticipate whether or not adverse impact might occur and to consider ways that minimize any exclusionary effects while preserving the ability to make valid inferences based on test scores. If adverse impact does occur, it is important

to demonstrate that the inferences made based on test scores are appropriate. By doing this, a company has the data to support the use of the test. U.S. case law and guidelines have clearly established that well-developed and validated tests can withstand legal scrutiny. Employers should have clear documentation regarding any tools they use in employment decision-making.

Practical constraints. Tests may not be the best choice if not many individuals are being considered in a particular employment decision, if the resources to properly administer the test are not available, or if the timing and logistics of the decision-making process preclude the use of an appropriate test.

The current decision-making process would not be improved upon by the addition of a test. Employers may believe they already have a quality decision-making process in place and a test would simply add costs and time with no gain in decision accuracy. Often, however, this belief has not been well-assessed, as organizations do not always track the information necessary to actually evaluate how well their employment decision-making processes are working. A proper evaluation of a decision-making process may reveal room for improvement, and often a test is a cost-effective and efficient way to improve the process.

Tests are useful decision making tools in employment contexts. Deciding whether a test is the right solution in a given situation may require professional advice from someone with knowledge of both testing and employment situations. Industrial-organizational psychologists may be helpful in such a situation.

2 How Many U.S. Companies Use Employment Tests?

Did you know that...*

41% of employers test job applicants in basic literacy and/or math skills.

34% of job applicants tested in 2000 lacked sufficient skills for the positions they sought.

68% of employers engage in various forms of job skill testing.

29% of employers use one or more forms of psychological measurement or assessment.

20% of employers use cognitive ability tests (for more information, see the article entitled [Types of Employment Tests](#)).

8% of employers use interest inventories.

14% of employers use some form of managerial assessments (for more information, see the articles entitled [Types of Employment Tests](#) and [Individual Psychological Assessment](#)).

13% of employers use personality tests (for more information, see the article entitled, [Types of Employment Tests](#)).

10% of employers use physical simulations of job tasks (for more information, see the article entitled, [Types of Employment Tests](#)).

****This information is based on survey data collected in 2001 from 1,627 HR managers who are members of the American Management Association. The survey accurately reflects the AMA membership base, and is therefore more reflective of larger organizations than all employers in the U.S. economy. Further, it is not reflective of test use outside the U.S. For more information on this survey, visit www.amanet.org/research/pdfs/bjp_2001.pdf.***

3 Establishing an Effective Employee Testing Program

There are many elements to establishing a testing program. This section outlines some of the more important elements. While establishing a program according to legal and professional standards does not guarantee that there will not be a legal charge or that an employer will be able to successfully defend the testing program, following legal and professional guidelines increases the probability of a successful defense. Many of these steps may require the services of a professional with training in test development and evaluation in employment settings.

- 1. Identify jobs where testing might be helpful.** The use of testing should be dependent on a careful analysis of the costs and benefits. Testing is often beneficial when there are high volumes of individuals to make decisions about, when there is high turnover, or when the consequences of making a poor decision are potentially very negative.

Managers should take time to carefully consider what the problem is before concluding testing is the solution. One of the biggest errors managers make is assuming their problem is a decision-making one, when, in fact, the problem is caused by poor training, poor supervision, or poor compensation. Once the manager defines the problem, he/she, with the help of a professional, can determine if a test will help or not.

Another consideration is assessing the size of the problem. If an organization makes decisions about only a few people a year, testing may not be cost effective, depending on the costs of poor performance, turnover and other negative decision outcomes for those few instances.

As noted earlier, testing can be helpful for improving employee performance, decreasing turnover, and decreasing costs associated with negative employee behaviors (e.g., theft, sabotage); a careful analysis will indicate whether testing is the right solution for a given situation.

- 2. Define job and organizational requirements clearly, completely, and accurately.** A common mistake employers make when hiring or promoting a job candidate is paying insufficient attention to what the requirements really are. Employers should gather systematic information on what are the knowledge, skills, abilities and other characteristics required for a given position (or entrance to a training program), and which are the most important to success. Not all key requirements will be easy to measure or should all be measured via testing; however, employers should ensure that any requirements that are assessed are important ones. In addition to job requirements, employers may go even further to define requirements as related to the organization's values and purpose.

3

- 3. Determine whether testing is useful for evaluating requirements.** Review the key job requirements and determine the best means of assessing these. This determination will be based on whether a requirement can easily be evaluated by a test, whether a test exists or must be developed for a requirement, whether a cost effective means of evaluation is available, whether available tests have demonstrated effectiveness, and other factors (see section on Information to Obtain When Creating or Purchasing a Test). This is where professional help can be of great value since I-O psychologists know the pros and cons of different kinds of tests or different purposes.
- 4. Select or develop a test (for more information about this topic, see the article entitled, Information to Obtain When Creating or Purchasing an Employment Test).** Obtain the names of test publishers and products from professional colleagues or employee testing professionals. Before purchasing a test, ask for information regarding the reliability and validity of the test. Reliability refers to the consistency of test results. Validity refers to whether the inferences made on the basis of a test score are correct. Also, ask for information on the development of the test and for information about the comparability of test results for different demographic groups. Evaluating the quality of this information is often difficult—while knowledgeable and experienced test publishers provide thorough and accurate information, others may attempt to pass off weak or inappropriate research and analyses as evidence of reliability and validity. Once again, competent professional evaluations of testing information can be helpful.

Rather than purchasing a test, employers may choose to develop their own test. This would be a good approach when an off-the-shelf product might not meet needs in the same way a customized tool would (e.g., using organization specific and job specific language in the test, assessing a requirement for which tests are not currently available, integrating with an existing testing system). Once again, the involvement of a professional industrial-organizational psychologist would be important to good test development.

- 5. Implement the test.** Make sure that the people who administer, score, and evaluate test results are appropriately trained. Depending on the type of test, such training might include issues related to standardizing administration conditions, scoring protocols, detecting cheating, what a score means, how to use test scores in decision-making, how to avoid rater biases, and many other things. Accurate record keeping of test scores and decisions made about individuals is typically a legal requirement, and is a necessity if one wishes to evaluate the effectiveness of the testing program.

Testing procedures should be consistent for all individuals for which the test is being used. Policies should be developed and articulated regarding issues such as proper administration conditions, retesting intervals, eligibility for testing, access to test scores, and other implementation issues.

3

Because tests may be the basis for decisions about individuals or may only be one piece of information used in making employment decisions, careful thought should go into how test results are used. For example, test scores might be used in a pass/fail manner or combined with other information. Test results might be used to set up score ranges indicating likelihood of individual success on the job. Once again, professional consultation can ensure that test results are used in the most effective manner.

- 6. Evaluate the testing program.** When possible, employers should attempt to gather information to evaluate the effectiveness of the testing program. This may be difficult to do with any accuracy if the test is used with only small numbers of individuals, the test is used in a highly restrictive manner (i.e., few individuals obtain a positive decision outcome after testing), or the outcome desired from testing is not easy to assess in a short time period. Consult with a professional regarding what kinds of record keeping might enable conducting an effective evaluation of the test.

4

Types of Employment Tests

Introduction

Hundreds of tests are available to help employers in making decisions. In the section on **Information to Consider When Creating or Purchasing a Test** there is a discussion of what one needs to consider in choosing what test to use. In this section, general types of tests are described and their general pros and cons related. Tests discussed in this section are ones for which some general evidence of validity has been provided. A test is valid if the inferences made based on the test score are accurate (e.g., if we are correct in concluding that how well the individual does on the test tells us how well he/she will perform on the job). Some tests, such as graphology and polygraphs, have little evidence of validity for employment decision-making purposes and thus are not discussed here. All of the tests discussed here have been demonstrated to relate to one or more critical employment outcomes.

Before deciding that a particular type of test is what you want, however, you should first establish a) does testing make sense (see section on **Employment Testing Overview**) and b) what it is you want to assess (see section on **Establishing an Effective Employee Testing Program**). The determination of what it is you want to measure with the test should precede the determination of how you are going to measure it.

4

I. Assessment Centers

Assessment centers can be designed to measure many different types of job related skills and abilities, but are often used to assess interpersonal skills, communication skills, planning and organizing, and analytical skills. The assessment center typically consists of exercises that reflect job content and types of problems faced on the job. For example, individuals might be evaluated on their ability to make a sales presentation or on their behavior in a simulated meeting. In addition to these simulation exercises, assessment centers often include other kinds of tests such as cognitive ability tests, personality inventories, and job knowledge tests. The assessment center typically uses multiple raters who are trained to observe, classify, and evaluate behaviors. At the end of the assessment center, the raters meet to make overall judgments about peoples' performance in the center.

Advantages:

- Have been demonstrated to produce valid inferences for a number of organizational outcomes (e.g., promotion rates).
- Can reduce business costs by identifying individuals for hiring, promotion or training who possess the needed skills and abilities.
- Can provide useful feedback to test takers regarding needed training and development.
- Focus more heavily on behavior demonstration than simply assessing characteristics.
- Use trained raters.
- Are typically less likely to differ in results by gender and race than other types of tests. other methods.

Disadvantages:

- Can be costly to create and administer.
- Require more labor (e.g., assessors, role-players, etc.) to administer than most other methods.
- Require more time to administer than most
- May be viewed positively by test takers who see the close relationship between the test and the job.
- Can be difficult to keep calibrated or standardized across time and locations.

2. Biographical Data

The content of biographical data instruments varies widely, and may include such areas as leadership, teamwork skills, specific job knowledge and specific skills (e.g., knowledge of certain software, specific mechanical tool use), interpersonal skills, extraversion, creativity, etc. Biographical data typically uses questions about education, training, work experience, and interests to predict success on the job. Some biographical data instruments also ask about an individual's attitudes, personal assessments of skills, and personality.

4

Advantages:

- Can be administered via paper and pencil or computerized methods easily to large numbers.
- Can be cost effective to administer.
- Have been demonstrated to produce valid inferences for a number of organizational outcomes (e.g., turnover, performance).
- Are typically less likely to differ in results by gender and race than other types of tests.
- Does not require skilled administrators.
- Can reduce business costs by identifying individuals for hiring, promotion or training who possess the needed skills and abilities.

Disadvantages:

- May lead to individuals responding in a way to create a positive decision outcome rather than how they really are (i.e., they may try to positively manage their impression or even fake their response).
- Do not always provide sufficient information for developmental feedback (i.e., individuals cannot change their past).
- Can be time-consuming to develop if not purchased off-the-shelf.

3. Cognitive Ability Tests

Cognitive ability tests typically use questions or problems to measure ability to learn quickly, logic, reasoning, reading comprehension and other enduring mental abilities that are fundamental to success in many different jobs. Cognitive ability tests assess a person's aptitude or potential to solve job-related problems by providing information about their mental abilities such as verbal or mathematical reasoning and perceptual abilities like speed in recognizing letters of the alphabet.

Advantages:

- Have been demonstrated to produce valid inferences for a number of organizational outcomes (e.g., performance, success in training).
- Have been demonstrated to predict job performance particularly for more complex jobs.
- Can be administered via paper and pencil or computerized methods easily to large numbers.
- Can be cost effective to administer.
- Does not typically require skilled administrators.
- Can reduce business costs by identifying individuals for hiring, promotion or training who possess the needed skills and abilities.
- Will not be influenced by test taker attempts to impression manage or fake responses.

Disadvantages:

- Are typically more likely to differ in results by gender and race than other types of tests.
- Can be time-consuming to develop if not purchased off-the-shelf.

4

4. Integrity Tests

Integrity tests assess attitudes and experiences related to a person's honesty, dependability, trustworthiness, reliability, and pro-social behavior. These tests typically ask direct questions about previous experiences related to ethics and integrity OR ask questions about preferences and interests from which inferences are drawn about future behavior in these areas. Integrity tests are used to identify individuals who are likely to engage in inappropriate, dishonest, and antisocial behavior at work.

Advantages:

- Have been demonstrated to produce valid inferences for a number of organizational outcomes (e.g., performance, inventory shrinkage difficulties in dealing with supervision).
- Can reduce business costs by identifying individuals who are less likely to be absent, or engage in other counterproductive behavior.
- Send the message to test takers that integrity is an important corporate value.
- Are typically less likely to differ in results by gender and race than other types of tests.
- Can be administered via paper and pencil or computerized methods easily to large numbers.
- Can be cost effective to administer.
- Does not require skilled administrators.

Disadvantages:

- May lead to individuals responding in a way to create a positive decision outcome rather than how they really are (i.e., they may try to positively manage their impression or even fake their response).
- May be disliked by test takers if questions are intrusive or seen as unrelated to the job.

5. Interviews

Interviews vary greatly in their content, but are often used to assess such things as interpersonal skills, communication skills, and teamwork skills, and can be used to assess job knowledge. Well-designed interviews typically use a standard set of questions to evaluate knowledge, skills, abilities, and other qualities required for the job. The interview is the most commonly used type of test. Employers generally conduct interviews either face-to-face or by phone. (For more information on this topic, see the article entitled, **Effective Employee Interviews**).

4

Advantages:

- Are expected and accepted by many job applicants.
- Provide an opportunity for a two-way exchange of information.
- Provide a measure of skills such as oral communication skills not measured via paper and pencil or computerized tools.
- Have been demonstrated to produce valid inferences for a number of organizational outcomes, if properly developed and administered (see article on **Effective Interviews**).
- Can reduce business costs by identifying individuals for hiring, promotion or training who possess the needed skills and abilities.
- Are typically less likely to differ in results by gender and race than other types of tests.

Disadvantages:

- May be affected by different kinds of rating errors and biases by interviewers.
- Are often more time-consuming to administer than paper and pencil or computerized tools.
- May be practically less useful when a large number of individuals must be evaluated because of administration time.
- Can be costly to train interviewers.
- May be difficult to keep interviewers calibrated and the interview process standardized.
- May lead to individuals responding in a way to create a positive decision outcome rather than how they really are (i.e., they may try to positively manage their impression or even fake their response).

6. Job Knowledge Tests

Job knowledge tests typically use multiple choice questions or essay type items to evaluate technical or professional expertise and knowledge required for specific jobs or professions. Examples of job knowledge tests include tests of basic accounting principles, A+/Net+ programming, and blueprint reading.

Advantages:

- Have been demonstrated to produce valid inferences for a number of organizational outcomes, such as job performance.
- Can reduce business costs by identifying individuals for hiring, promotion or training who possess the needed skills and abilities.
- Are typically less likely to differ in results by gender and race than other types of tests.
- May be viewed positively by test takers who see the close relationship between the test and the job.
- Will not be influenced by test taker attempts to impression manage or fake responses.
- Can provide useful feedback to test takers regarding needed training and development.

Disadvantages:

- May require frequent updates to ensure test is current with the job.
- May be inappropriate for jobs where knowledge may be obtained via a short training period.
- Can be costly and time-consuming to develop, unless purchased off-the-shelf.

4

7. Personality Tests

Some commonly measured personality traits in work settings are extraversion, conscientiousness, openness to new experiences, optimism, agreeableness, service orientation, stress tolerance, emotional stability, and initiative or proactivity. Personality tests typically measure traits related to behavior at work, interpersonal interactions, and satisfaction with different aspects of work. Personality tests are often used to assess whether individuals have the potential to be successful in jobs where performance requires a great deal of interpersonal interaction or work in team settings.

Advantages:

- Have been demonstrated to produce valid inferences for a number of organizational outcomes.
- Can reduce business costs by identifying individuals for hiring, promotion or training who possess the needed skills and abilities.
- Are typically less likely to differ in results by gender and race than other types of tests.
- Can be administered via paper and pencil or computerized methods easily to large numbers.
- Can be cost effective to administer.
- Does not require skilled administrators.

Disadvantages:

- May contain questions that do not appear job related or seem intrusive if not well developed.
- May lead to individuals responding in a way to create a positive decision outcome rather than how they really are (i.e., they may try to positively manage their impression or even fake their response).
- May be problematic for use in employee selection if the test is one used to diagnose medical conditions (i.e., mental disorders) rather than simply to assess work-related personality traits.

8. Physical Ability Tests

Physical ability tests typically use tasks or exercises that require physical ability to perform. These tests typically measure physical attributes and capabilities, such as strength, balance, and speed.

Advantages:

- Have been demonstrated to produce valid inferences regarding performance of physically demanding tasks.
- Can identify applicants who are physically unable to perform essential job functions.
- Can reduce business costs by identifying individuals for hiring, promotion or training who possess the needed skills and abilities, by minimizing the risk of physical injury to employees and others on the job, and by decreasing disability/medical, insurance, and workers compensation costs.
- Will not be influenced by test taker attempts to impression manage or fake responses.

Disadvantages:

- Are typically more likely to differ in results by gender than other types of tests.
- May be problematic for use in employee selection if the test is one used to diagnose medical conditions (i.e., a physical disability) rather than simply to assess ability to perform a particular job-related task.
- Can be expensive to purchase equipment and administer.
- May be time consuming to administer.
- May be inappropriate or difficult to administer in typical employment offices.

4

9. Work Samples and Simulations

These tests typically focus on measuring specific job skills or job knowledge, but can also assess more general skills such as organizational skill, analytic skills, and interpersonal skills. Work samples and simulations typically require performance of tasks that are the same or similar to those performed on the job to assess their level of skill or competence. For example, work samples might involve installing a telephone line, creating a document in Word, or tuning an engine. (For more information on this topic, see the article entitled, Work Samples and Simulations).

Advantages:

- Have been demonstrated to produce valid inferences regarding ability to perform the job.
- Can reduce business costs by identifying individuals for hiring, promotion or training who possess the needed skills and abilities.
- Are less likely to differ in results by gender and race than other types of tests (depends on particular skills being assessed).
- May be more accepted by test takers due to the obvious link between the test and the job.
- Less likely to be influenced by test taker attempts to impression manage or fake responses.
- Can be used to provide specific developmental feedback.
- Can provide test takers with a realistic preview of the job and the organization.

Disadvantages:

- Does not assess aptitude to perform more complex tasks that may be encountered on the job.
- May not assess the ability to learn new tasks quickly.
- Often not conducive to group administration.
- May require some level of job knowledge and therefore may be inappropriate for jobs where knowledge may be obtained via a short training period.
- May be difficult to keep updated.
- May be expensive to administer.
- May be time consuming to develop and to administer.

5 Types of Test Item Formats

Introduction

Just as there are several types of tests available to help employers make employment decisions, there are also several types of test formats. In this section, the pros and cons of general types of test item formats are described. Also, some general guidelines for using different types of test item formats are provided.

Before deciding on a particular type of test format, you should first establish a) does testing make sense (see section on **Employment Testing Overview**) and b) what it is you want to assess (see section on **Establishing an Effective Employee Testing Program**). The determination of what it is you want to measure with the test should precede the determination of how you are going to measure it.

Pros and Cons of Multiple Choice Test Items:

PROS:

- Can be used to test many levels of learning.
- Can be used to test a person's ability to integrate information.
- Can be used to diagnose a person's difficulty with certain concepts.
- Can provide test takers with feedback about why distractors were wrong and why correct answers were right.
- Can ask more questions, greater coverage of material.
- Can cover a wide range of difficulty levels.
- Usually requires less time for test takers to answer.
- Usually easily scored and graded.

CONS:

- Test takers may perceive questions to be tricky or too picky.
- Difficult to test attitudes towards learning because correct responses can be easily faked.
- Does not allow test takers to demonstrate knowledge beyond the options provided.
- Requires a great deal of time to construct effective multiple choice questions, especially ones that test higher levels of learning.
- Encourages guessing because one option is always right.
- Test takers may misinterpret questions.

Pros and Cons of True-False Test Items

PROS:

- Can ask more questions for greater coverage of material.
- Can cover a wide range of difficulty levels.
- Usually requires less time for test takers to answer.
- Usually easily graded and scored.

CONS:

- Does not allow test takers to demonstrate broad range of knowledge
- Is difficult to construct effective true-false items that test higher levels of learning
- Encourages guessing due to 50/50 chance of being correct
- Is easily faked, difficult to test attitudes toward learning

Pros and Cons of Essay Test Items

PROS:

- Can test complex learning objectives.
- Can test processes used to answer the question such as the ability to integrate ideas and synthesize information.
- Requires use of writing skills, correct spelling, and grammar.
- Can provide a more realistic and generalizable task for test.
- Usually takes less time to construct.
- Is more difficult for test takers to guess correct answer.

CONS:

- Usually takes more time to answer.
- Can be unreliable in assessing the entire content of a course or topic area.
- Essay answers are often written poorly because test takers may not have time to organize and proofread answers.
- Is typically graded or scored more subjectively; non-test related information may influence scoring process.
- Requires special effort to be graded in an objective manner.
- Requires more time to grade or score.

Guidelines for Using Multiple Choice or True-False Test Items

It is generally best to use multiple-choice or true-false items when:

1. You want to test the breadth of learning because more material can be covered with this format.
2. You want to test different levels of learning.
3. You have little time for scoring.
4. You are not interested in evaluating how well a test taker can formulate a correct answer.
5. You have a clear idea of which material is important and which material is less important.
6. You have a large number of test takers.

5

Guidelines for Using Essay Test Items

It is generally best to use essay items when:

1. You want to evaluate a person's ability to formulate a correct answer.
2. You want to assess a people's ability to express themselves in writing and writing is an important aspect of the job.
3. You have time to score the essay items thoroughly.
4. You feel more confident about your ability to read written answers critically than to construct effective multiple-choice items.
5. You want to test a person's ability to apply concepts and information to a new situation.
6. You have a clear idea of the most important information and concepts that should be tested.

6

Effective Interviews

While most organizations use interviews in employment decision making, many have not maximized the effectiveness of their interviewing process. In the section on Types of Employment Tests, the advantages and disadvantages of interviews were presented. In this section, we discuss what distinguishes an effective interview from an ineffective one. Research has been fairly conclusive in showing that structuring an interview—making sure that the characteristics to be evaluated are clearly identified, that all interviewers ask the same questions of all candidates, that interviewers are well trained, and that interviewee performance is evaluated using well-developed rating scales—leads to a more effective interview process.

The following tips can help managers create an effective interview:

1. *Identify the candidate characteristics that should be assessed during the interview.*

Employers must consider a number of factors in deciding what candidate characteristics will be evaluated in the interview. Questions to be considered include: Is this a knowledge, skill, or characteristic that is important to success on the job or to some outcome of interest (e.g., low turnover)? Is the interview the best way to assess this important knowledge skill or ability? How much overlap would be best between the interview and other tests used in the decision-making process?

2. *Develop interview questions.*

Once the areas to be evaluated are identified, interview questions should be created that will be used with all candidates for a particular position. Past behavior is one of the best predictors of future behavior. Interview questions should be designed so that candidates describe things they actually did or said in a previous situation and the outcome of their actions. This information often predicts very well how candidates are likely to respond to a similar situation in the future. While questions about hypothetical situations (what would you do if . . .) can be useful also, these questions need to be worded and used carefully within any interview to get the maximum benefit from them. For example, asking an individual about whether he/she would apply a skill in a particular situation can result in his/her giving an obvious, socially desirable response.

6

3. Plan likely probes and follow-up questions.

Organizations may find that an interview process that was designed to be one in which all interviewees are treated the same and asked the same questions becomes one in which people are asked dramatically different things because interviewers vary widely in what they do in follow-up questions or probing. If interviewers are not allowed to probe, however, often key information is not elicited. Thus, the interview should be designed to give interviewers the freedom to ask follow-up questions but also to guide them in the types of follow-up questions that would be most appropriate for the given interview structure. A suggested list of possible probes can accompany the list of interview questions, and/or training in effective probing be provided to interviewers.

4. Evaluate responses using anchored scales.

A systematic evaluation of individuals' responses to interview questions is helpful for several reasons. It allows for a comparison across candidates who are often interviewed by different individuals or even by the same individual across a wide time span. It requires the interviewer to evaluate the candidate on the job-relevant characteristics identified as important, not on any idiosyncratic set of criteria. Standardization in rating scales provides documentation that all candidates were evaluated on the same basis.

5. Train interviewers.

Despite the beliefs of many individuals that they are "good judges" of others, interviewer training has been demonstrated to be effective in improving the judgments of interviewers, in ensuring that all candidates are treated similarly, and in calibrating interviewers with one another.

6. Understand the legal parameters.

Using business-related job requirements as the foundation for creating interview questions usually means the interview will be in compliance with the law. In deciding which questions to ask in an interview, understanding the legal requirements for selection will help the interviewer avoid asking inappropriate questions. For example in the U.S., laws such as Title VII of the Civil Rights Act (1964), the Americans With Disabilities Act of 1990 and the Age Discrimination in Employment Act of 1967 outline several areas of concern with regard to employment discrimination. These characteristics include race, color, gender, religion, national origin, age (over 40), and disabilities. Generally, questions directly about these characteristics will be problematic, but questions that ask about a job requirement may relate to a characteristic and be acceptable. For example, one cannot ask "Are you a member of a religion that holds services on Saturday?" but one can ask, "This job requires overtime work on the weekends. Can you work on Saturdays and Sundays?" In other countries, and in specific jurisdictions within the U.S., other issues may be of concern. Contact your local governmental agency for lists of questions to avoid.

6

7. Use the interview to provide a realistic job preview.

Interviews are not just opportunities to learn more about candidates; they are also opportunities for managers to help candidates learn more about the job. When candidates have a realistic understanding of their job, their expectations are more likely to be met. When a job fails to live up to an individual's expectations, he/she is more likely to be dissatisfied and ultimately leave the position. When encountering tight labor markets, managers may be inclined to only relay the positive aspects of a job to candidates. Although this approach may result in short-term success, ultimately more time is spent hiring as candidates continually leave the position.

8. Use the interview as a selling opportunity.

To increase the chances of hiring a good candidate, use the interview experience to sell the job and the company. Prior to the interview, talk to employees in the company and find out what they like best about the organization. Then, when talking to candidates about what they are looking for in a work experience, relay how the job and the organization can meet their needs. To make the most of this selling opportunity, think about the interview experience from the candidates' perspective. A candidate's experience in the interview process affects his/her opinions about the organization. Being treated professionally and talking with a well-prepared interviewer creates a positive impression and experience.

Developing an effective interview process is challenging. The Society for Industrial and Organization Psychology as well as the psychology and business departments of your local college or university are excellent resources for consultants that can answer any questions about interview development.

7 Work Samples and Simulations

What better way to assess an applicant's job potential than to actually have them complete a portion of the work they will be doing? In addition to more traditional testing methods (e.g., cognitive tests, personality questionnaires, etc.), many organizations include **work samples** or **simulations** in their decision-making processes.

Such exercises are carefully constructed to simulate on-the-job situations. This type of exercise can range from something as simple as having the candidate design and deliver a presentation (to assess presentation skills for a training position) to having candidates complete an elaborate, highly structured computerized simulation (to be used to assess one's ability to operate expensive, potentially dangerous equipment). Two common simulations are the In-Basket exercise and the Role-Play.

In-Basket Exercises

In-Basket exercises are typically designed to simulate the administrative tasks of a job. In the typical In-Basket, the test taker is given background information on either the actual organization or a fictitious company and is asked to assume a specific role in the organization. The test taker is often asked to respond to letters, memos, e-mails, requests, personnel issues, and so forth, in a given amount of time. A time limit may be imposed in order to simulate the time pressure experienced in many jobs. In-Basket exercises are usually designed to assess the candidate's ability to manage multiple tasks, prioritize and delegate work, and analyze information quickly. In-Basket exercises may be administered via paper-and-pencil methods or via computer. Scoring of In-Baskets varies considerably, with some more easily scored via computerized templates and others requiring careful review by a trained evaluator.

Role-Plays

Role-Play exercises are designed to simulate the interpersonal challenges faced when working with others. In the typical role-play, the candidate is given background information regarding the scenario and asked to play a particular role (e.g., team leader, customer service representative). During the exercise, he or she interacts directly with a trained role-player (actor). This actor often plays the role of a subordinate, coworker, or customer and responds to the candidate according to a script. Role-Play exercises are usually designed to assess the candidate's communication and interpersonal skills. Performance may be observed by a trained evaluator, or may be videotaped and evaluated at a later time.

7

Keys to the Effective Use of Simulations

There are many advantages to an employer in using work samples and simulations (see **Types of Employee Tests** section for a listing). Decisions to use a work sample or simulation should be made with these considerations in mind. Regardless of the type of exercise used, all simulations will be most effective if they:

1. Are based on thorough and accurate job information.
2. Are constructed with a consideration of guidelines for quality test development.
3. Are highly structured, so that all individuals are given the same opportunities, and all are evaluated on the same basis.
4. Include some type of standardized rating scale to score responses.
5. Include multiple raters where appropriate and possible.

8

Individual Psychological Assessment

Hiring, promoting, and developing the right people for specific jobs represent some of the most important decisions that companies make. Many organizations use individual psychological assessment to help them make such decisions. Individual psychological assessment or “individual assessment” is a phrase used by many people. Often, that phrase means different things to different people. The following questions and answers are presented to clarify this phrase.

1. What is individual psychological assessment?

Individual psychological assessment is a tool used to help organizations make decisions about hiring, promotion, and development. A typical individual psychological assessment consists of professionally developed and validated measures of personality, leadership style, and cognitive abilities among other things. The process often includes an interview.

Ideally, the tests and measures used in an individual assessment have been validated for the current positions in accordance with legal and professional guidelines (see APA Standards and SIOP Principles). While the small number of individuals evaluated for certain positions may make obtaining certain kinds of validation evidence more difficult, other types of evidence should be accumulated to support the use of individual assessments.

2. How much time does a typical individual psychological assessment take?

Depending on factors such as the length of an interview and the number of tests and simulations that are included in the assessment, the length of the process can vary. The level of the position (e.g., entry level, executive, etc.) can also affect the amount of time an organization chooses to spend on the individual psychological assessment process. The process varies from less than an hour to two days.

3. Who are the primary users of individual psychological assessment?

Organizations of all sizes may include individual assessment as supplemental information for making hiring, promotion, and development decisions.

4. Who is qualified to conduct individual psychological assessments?

There are different roles that are part of the individual assessment process, and different people can perform these roles. For example, paper-and-pencil tests vary in their requirements for administration, scoring, and interpretation (i.e., some can be administered and scored by trained administrative staff members while others require a background in testing for administration, scoring and interpretation). Typically, for an individual psychological assessment, the assumption is that a qualified psychologist combines and interprets the data from the various assessment tools to make an overall evaluation of an individual.

8

5. How is individual psychological assessment information generally used?

Individual psychological assessment information is usually intended to supplement data on a candidate obtained from other sources (e.g., interviews, references, and prior work experience). The results of the individual psychological assessment are often considered as information to be explored further with the candidate and with his or her references.

6. When is individual psychological assessment typically used?

Due to cost and time, the individual assessment process is usually implemented as a final hurdle when the candidate pool has been narrowed down to the top candidates. Many organizations assess at least two individuals for comparison purposes. Assessing at least two candidates can provide a back up in case the preferred candidate declines the job offer. Comparing two or three candidates with one another also provides the company with the additional advantage of determining relative fit with the culture and the potential for coaching and development. Also, because of time and cost considerations, individual assessment is more often used for high level or critical positions in the organizations, where the costs of wrong employment decisions are high.

7. How much does a typical individual psychological assessment cost?

Fees for individual assessment are typically based on several factors, including the number of tests used, the length of the assessment, the professional training and experience of the psychologist, the reputation of the consulting firm, and its geographic location. Companies are more likely to spend more on higher level positions where the stakes are high.

8. What are the advantages and disadvantages of individual psychological assessment relative to other employee screening and evaluation methods?

The primary benefit of individual psychological assessment is that it provides additional data on a candidate that is not easily obtained through using only one or two assessment tools. Another benefit of individual psychological assessment is that it can provide a candidate with detailed, developmental feedback for the new position. However, developmental feedback is often optional in individual assessment and is not always provided. The data can also help the candidate's new manager gain a better understanding of how best to coach and mentor his or her new direct report. However, depending on the organization's policy and the appropriate use of the assessment information, the data may or may not be shared with the supervisor.

The primary disadvantages are the time and expense associated with the assessment. Also, the process may alienate or "turn off" some candidates who do not understand the process or who do not see it as relevant to the decision.

9

Information to Consider When Creating or Purchasing an Employment Test

Employers have two choices when implementing an employment test. They can either purchase a test or create their own test. There are many employee testing products and services on the market today. When deciding to purchase a test, managers may experience information overload when reviewing information on testing products and services. Most test publishers provide a technical summary or manual that describes the most important qualities and characteristics for any given test. The technical manual should provide information on most, if not all, of the factors to consider before purchasing a test. Obtaining professional help in interpreting testing information is often necessary. While most managers will not have the time, resources, or background to engage in test development, this material is provided to assist the manager in planning and reviewing work done by professionals.

Creating a test is a complex and time-consuming process, so experts either inside or outside the organization should be involved if a decision is made to create rather than purchase a test. Employers develop their own tests for a variety of reasons such as cost effectiveness, test security concerns, company culture, position uniqueness, and other factors. The need to develop a test is often based on a cost/benefit analysis. This analysis addresses the cost to develop and use one's own test compared to buying someone else's test and paying to use it over the life of the instrument.

If purchasing a test, obtaining information on these factors is important. If creating a test, developing documentation related to these factors is important. These factors apply regardless of what you plan to measure with the test, what type of test you are considering, or what mode of administration (computer vs. paper-and-pencil vs. performance) you are planning.

- 1. Test development information.** The research and development that went into creating the instrument should be documented. What was the theory or experience on which the test was based? Was the test developed on people that are similar to this organization's applicants or employees? What was the process used to develop the test? At the very least, this background information on the test is important because it provides information on the logic, care, and thoroughness by which the test was developed.
- 2. Reliability.** Reliability refers the consistency of test results. There are several ways to assess the reliability of a test, and some are more appropriate for certain situations (e.g., when multiple raters or evaluators are involved; if one wishes to know about stability of results over time). Experienced and knowledgeable test publishers have information on the reliability of their testing products.

- 3. Validity.** Validity refers to the accuracy of the inferences made based on test results (e.g., how accurate is it to say that a higher test score indicates that a person is more likely to be a better performer). There are many forms of validity evidence. For example, evidence might consist of showing a relationship between test scores and some outcome of interest (e.g., supervisory ratings of job performance, average monthly sales, turnover). Evidence might consist of documentation of links between the content of the test and the requirements of the job. Evidence might include showing that the test relates to other measures of the same thing. Experienced and knowledgeable test publishers have (and are happy to provide) information on the validity of their testing products. Judgments regarding what types of validity evidence are appropriate for a given test depend on a number of factors, and these are outlined in *The Standards for Educational and Psychological Testing* (www.apa.org/science/standards.html). Trained professionals can help interpret whether the evidence supporting the particular inferences an employer wishes to make with a test is sufficient.
- 4. Test bias.** Test developers should provide evidence that the test does not contain bias on the basis of race or sex, that is, that the test is related to outcomes in a similar manner for all individuals. This does not mean that the test will have similar results for different groups of people, but that it is not a biased indicator of an outcome of interest. For example, in a typical employment decision context, more women than men will score low on a test of upper body strength, but the test would not be considered biased if women and men with similar scores achieved similar performance on the job.
- 5. Information on administration.** The documentation should include a description of all materials required for administration (e.g., test booklets, answer sheets, scoring keys, etc.) and administration instructions. Instructions should discuss issues such as standardization of testing conditions (e.g., noise, lighting, time) and how to avoid nonstandardized administrations. Qualifications for administering should also be clearly stated.
- 6. Data for test interpretation.** Test scores cannot be interpreted in isolation, but are interpreted in light of other information. For example, whether a test score is considered good or poor may depend on the distribution of scores of a comparison group. This comparison group is typically referred to as a norm group. The test publisher should provide information about the different norm groups that are available for the test being considered. Ideally, one uses a norm group that is similar to the group of people that are in the position for which testing is being used. There are other ways to interpret test results including expectancy charts and cut scores, which are developed based on information about how the score relates to outcomes of interest. Information should be made available on data that can aid in appropriate test score interpretation.
- 7. Scoring options.** Determine what scoring and test reporting needs are to determine whether the test has appropriate options. Some tests can be scored on-site, either by hand or by machine. Other tests require that an employer call, mail, or fax the test results to the test publisher for scoring. Qualifications for scoring should be explicit.

9

8. **Ongoing research/refinement of the test.** The test publisher should indicate when the test was developed and when the test was last updated. Test publishers often update their tests to comply with new legal requirements or to reflect changes in vocabulary or terminology.
9. **Time requirements.** Some tests have time limits while others provide the test taker with unlimited time to take the test. How time limits were determined and why they are necessary should be documented. For tests that have time limits, greater administrator training may be needed.
10. **Credentials and experience.** The educational background and work experience of the persons who developed the test should be documented, as well as references that can speak to the capabilities and experience of the test developer or vendor. Some tests require the test administrator or individuals interpreting test scores to have certain credentials (e.g., MA, PhD) that reflect coursework in statistics, test interpretation, or test development and validation.
11. **Cost.** Direct costs of the test usually include test booklets, answer sheets, and a test administrator's manual. Hand-scored tests usually include a scoring template as well. Computer-based testing generally includes software that is valid for a prepaid number of uses. Testing fees usually need to be considered as an ongoing expense, since few test publishers will license a test for unlimited usage. This situation often leads organizations to create their own tests. However, remember that part of the cost of the test is for the substantial investment that the publisher made in researching and developing a high quality measure. Creating a high-quality and effective test requires time, money, and people to research and develop, revise, and validate the tests. In many instances, the more cost effective approach is to purchase a test from a test publisher rather than to create a test.

Choosing or developing a test is a challenge, and this list just mentions some of the information you should obtain. A comprehensive source of the information you should consider in testing is provided by the ***Standards for Educational and Psychological Testing*** (www.apa.org/science/standards.html). You might also want to check out, www.onetcenter.org, for ***“Testing and Assessment: An Employer’s Guide to Good Practices.”*** This guide helps managers and others understand and use employment testing and assessment practices.

10 Seven Questions to Ask a Vendor Before Purchasing a Test

You've decided that testing is a viable and useful option for your organization. You also know what knowledge, skills, abilities, or other characteristics you want to assess. Now what? How do you find a good quality test or tests that measures what you want to measure?

Thousands of employment testing products and services exist in the marketplace. They vary in many ways including cost, time to administer, and overall quality. When deciding to purchase a test, managers typically experience information overload. Obtaining professional help in interpreting testing information is often necessary. While most managers will not have the time, resources, or background to engage in test development, this material is provided to assist the manager in planning and reviewing work done by professionals.

To help you make a good decision related to employment testing, ask the following seven questions before purchasing a test. If a test publisher or vendor cannot or will not answer the following questions, you should be wary. Knowledgeable and experienced test publishers typically welcome a client's detailed interest in their product. Test publishers and vendors also want to provide whatever information is necessary to help an organization make a good "employment testing" related decision.

1. What does the test measure?

Look for a clear and concise answer to this question. It is difficult to develop a good quality test without a clear definition of what is to be measured. Also, consider how the answer compares with what needs to be assessed in your organization. Be wary of a vendor who describes their test as measuring many seemingly unrelated factors. Also be wary of tests that consist of a limited number of items—for example a 25-item test that purports to measure 10 different factors. Measuring a characteristic or quality consistently and accurately typically requires more than just a couple of items.

2. What research and process was used to develop the test?

What was the theory or experience on which the test was based? Was the test developed on people that are similar to your organization's applicants or employees? What was the process used to develop the test? At the very least, this background information on the test is important because it provides information on the logic, care, and thoroughness by which the test was developed.

10 3. **What experience and/or education do you have that qualifies you to develop and/or sell this test?**

The educational background and work experience of the persons who developed the test is important, as well as references that can speak to the capabilities and experience of the test developer or vendor. To have confidence in the test and in the event of a legal challenge, you want test developers or vendors who have education and/or experience related to the specific content of the test and related to test development and validation. Also, some tests require the test administrator or individuals interpreting test scores to have certain credentials (e.g., MA, PhD) that reflect coursework in statistics, test interpretation, or test development and validation.

4. **What evidence do you have related to the reliability of this test?**

Reliability refers to the consistency of test results. There are several ways to assess the reliability of a test, and some are more appropriate for certain situations (e.g., when multiple raters or evaluators are involved; if one wishes to know about stability of results over time). Experienced and knowledgeable test publishers and vendors have information on the reliability of their testing products. For more detailed information on how to assess the reliability of a test, check out ***Testing and Assessment: An Employer's Guide to Good Practices*** at www.onetcenter.org.

5. **What evidence do you have related to the validity of this test?**

Validity refers to the accuracy of the inferences made based on test results (e.g., how accurate is it to say that a higher test score indicates that a person is more likely to be a better performer). Knowledgeable and experienced test publishers typically have many forms of validity evidence. For example, they may have evidence that shows a relationship between test scores and some outcome of interest (e.g., supervisory ratings of job performance, average monthly sales, turnover). They might also have evidence that documents a link between the content of the test and the requirements of the job. Other evidence might include showing how the test relates to other measures of the same thing. Experienced and knowledgeable test publishers have (and are happy to provide) information on the validity of their testing products. Judgments regarding what types of validity evidence are appropriate for a given test depend on a number of factors, and these are outlined in ***The Standards for Educational and Psychological Testing*** (www.apa.org/science/standards.html). For more detailed information on how to assess the validity evidence associated with a particular test, check out ***Testing and Assessment: An Employer's Guide to Good Practices*** at www.onetcenter.org.

10 6. What evidence do you have that demonstrates the lack of bias or discrimination of your test?

Look for evidence that the test does not contain bias on the basis of race or sex, that is, that the test is related to outcomes in a similar manner for all individuals. This statement does not necessarily mean that the test will have similar results for different groups of people. This statement does mean that the test is not a biased indicator of an outcome of interest. For example, in a typical employment decision context, more women than men will score low on a test of upper body strength. The test, however, would not be considered biased if women and men with similar scores achieved similar performance on the job.

7. What data do you have that will help me interpret test scores in my organization?

You cannot interpret test scores by themselves. Whether a test score is considered good or poor may depend on the distribution of scores of a comparison group. This comparison group is typically referred to as a norm group. The test publisher should provide information about the different norm groups that are available for the test being considered. Ideally, you want to use a norm group that is similar to the group of people that are in the position for which testing is being used. There are other ways to interpret test results including expectancy charts and cut scores, which are developed based on information about how the score relates to outcomes of interest. Information should be made available on data that can aid in appropriate test score interpretation.

Using Tests Effectively in the Public Sector

There are far more similarities than differences when it comes to effective testing in the public versus private sector. Nevertheless, there are some unique differences in employment testing for hiring and promotion purposes in the public sector. In particular, public sector testing occurs in a public arena, and therefore often requires greater openness regarding the testing process and greater public availability of information regarding individual test performance, test score use, test development, and other aspects of the testing process.

Merit Systems

U.S. government agencies often base their hiring and promotional practices on a “merit system,” which uses a competitive examination process to assess the level or degree of qualification of candidates. These examinations may consist of one or more tests, such as multiple-choice tests, essays, interviews, physical ability tests, and performance tests.

Some tests may be evaluated on a pass/fail basis. One or more tests, however, must be scored on a more discrete basis, often percentage scores, to permit ranking of candidates. If more than one test is used, the scores can be weighted, and then combined to result in an overall score. Once candidates’ final numeric examination scores are computed, they are placed on an “eligible list” in descending rank order based on this score and are then considered in this order for hiring or promotion. Different agencies use different “rules” for designating those candidates who are eligible for final hiring/promotion. One of the more common rules is to allow consideration of only the three highest scoring candidates. Some agencies permit consideration of only the highest scoring candidate. And, at the other extreme, some agencies apply a “rule of the list” wherein any candidate on the rank ordered list may be considered.

Civil Service Examinations

The civil service examination is the primary mechanism for selection into and promotion up through the ranks of many U.S. government agencies. This results in agencies’ conducting a large number of highly structured multi-faceted examinations. Because many individuals are attracted to this type of employment setting, and because many agencies conduct extensive outreach recruitment, large candidate groups are often the norm. This situation routinely presents considerable challenges in constructing and administering rigorous, comprehensive examinations for jobs to which hundreds or even thousands of candidates apply.

Validation Strategy

Public agencies tend to rely on content-oriented test construction to demonstrate the job-relatedness of employment tests. There are several reasons for this tendency, including:

1. They follow a rigid job classification system, with different jobs being specifically and narrowly defined. Job descriptions or “class specifications” serve as a good starting point for constructing a more content-oriented test.
2. They often prefer to hire employees fully proficient in a given type of work when possible, making a content validation strategy particularly applicable. “Relative merit” is often defined in terms of “to do” the job. Defining relative merit in this manner reflects the preference for a high level of content knowledge/skillful performance among those hired.
3. Public sector employers may be required to have a relatively open system that allows for public scrutiny of processes, including candidate “protest” of examinations; and content-oriented tests, which usually exhibit a high degree of face validity that makes them more acceptable to laypersons, are less likely to attract criticism.
4. Resources are often limited among public sector employers, so their ability to conduct more elaborate test validation research (or even to purchase tests) is restricted. The narrowness of job classifications also serves to deter conducting criterion-validation studies because most jobs have too few incumbents to yield consistent results statistically.

12 Employment Testing Websites

More information about supporting the effectiveness of tests for employment decision-making purposes can be found in SIOP's Principles for the Validation of Employee Selection Procedures (link to spot on our site where it will be). In addition to the Society for Industrial and Organizational Psychology, visit the following Web sites for more information about employment testing.

Locate a Test/Test Reviews: *For more information about types of commercially available tests, test publishers, and reviews of various tests, contact the following organizations.*

1. Association of Test Publishers (www.testpublishers.org)

The Association of Test Publishers is a non-profit organization representing publishers and providers of tests and assessment tools and/or services related to assessment, selection, certification, and educational and clinical issues. This site has a useful section on testing called, Frequently Asked Questions.

2. Buros Institute of Mental Measurement (www.unl.edu/buros/)

Buros Institute of Mental Measurements has supported advancements in test practice since 1938. The Institute publishes the Mental Measurements Yearbook and Tests in Print series that includes reviews of commercially available tests.

Locate Information on Employment Testing: *For more information about legal and government requirements for employment testing as well as testing best practices in organizations, contact the following organizations.*

1. American Psychological Association's Testing Page (www.apa.org/)

The American Psychological Association produces a number of documents of particular importance to test development and use. These include the Standards for Educational and Psychological Testing, the Rights and Responsibilities of Test Takers and Test User Qualifications. They also have several Web pages that are useful for employers considering testing:

<http://www.apa.org/science/faq-findtests.html> This Web site provides answers to frequently asked questions about tests. In particular, it provides information on how to locate a test for a given characteristic and information on volumes that reference all sorts of tests.

<http://www.apa.org/science/standards.html> This page provides information on how to order The Standards for Educational and Psychological Testing, a comprehensive document that indicates what test users should be aware of and do when using tests.

12

2. Department of Labor—Employment and Training Administration (www.doleta.gov)

The Employment and Training Administration arm of the Department of Labor seeks to build up labor market through workforce training and the placement of workers. Employers will find information on several employment related areas including tax incentives, other hiring incentives, and information on how to find and train employees.

3. Equal Employment Opportunity Commission (www.eeoc.gov)

Created by the Civil Rights Act of 1964, the Equal Employment Opportunity Commission (EEOC) is responsible for enforcing the employment-related provisions of the act. This organization is the major government agency involved with employment discrimination. The EEOC initiates investigations, responds to complaints, and develops guidelines to enforce various laws.

4. International Personnel Management Association – Assessment Council (www.ipmaac.org)

The International Personnel Management Association’s Assessment Council (IPMAAC) is the leading organization of applied personnel assessment professionals. With more than 600 members, the organization provides a forum for sharing interests and expertise in the development and effective use of personnel assessment methods in organizations.

5. O*NET (www.onetcenter.org)

O*NET, the Occupational Information Network, is a comprehensive database of worker attributes and job characteristics. As the replacement for the Dictionary of Occupational Titles, O*NET is the nation’s primary source of occupational information. Check out this site for “Testing and Assessment: An Employer’s Guide to Good Practices.” This guide helps managers and others understand and use employment testing and assessment practices.

6. Society for Human Resource Management (www.shrm.org)

This site represents the national organization of human resource and personnel professionals within public and private organizations. It includes up-to-date HR-related news items and commentary and a long list of links to other HR-related sites.

7. Society of Industrial and Organizational Psychology (www.siop.org)

This site represents the national organization of industrial and organizational psychologists within public and private organizations. It includes accurate and objective information on workplace topics. SIOP also provides a service, the Consultant Locator System, that helps people find industrial and organizational psychologists and consulting firms that can assist them with various workplace issues.