

# Employee selection



Beginning in the 1970s, two developments dramatically changed in employee selections. First, the development of meta-analysis, arguably one of the most influential methodological developments in recent decades, made it possible to cumulate quantitatively the results of large numbers of small-scale studies, resulting in a quasi-massive-scale study. Second, the results of large-scale studies of military personnel and others also became available. The results of both kinds of studies provided strong evidence of remarkably general validity for cognitive ability tests for selection across a broad range of jobs. Given this state of affairs, it is not surprising that some have argued for near universal use of cognitive ability tests as the primary selection tool.

In addition to the positive results from meta-analytic and large-scale predictive-validity studies, cognitive ability tests are remarkably practical. After 85 years of research, cognitive ability tests are among the most reliable measures available to social scientists. Also, unlike selection tools such as checking references or evaluating prior performance, cognitive ability tests can be given to individuals who are new to the job market. Despite these strengths, others have argued that it is important to look beyond general cognitive ability if one is to understand why people achieve to the extent that they do on the job. The most important issue in HR selection testing is determining a test's validity.

The actual definition of validity can vary depending on the circumstances, the specific tools used, and the application. For most selection purposes, however, a selection test is valid if the characteristic(s) it is measuring is related to the requirements and/or some important aspect test is valid, and a

test is valid if there is a link between the test score and job performance. The degree to which an employment selection test has validity tells the testing entity what it can conclude or predict about someone's job performance from his or her test scores. A test's validity is established for a specific purpose, and it may not be valid for purposes other than those that it has been validated to measure. Criterion-related validity is the correlation or other statistical relationship between selection test score (the predictor) and job performance (the criterion).

If those who score low on a test also perform poorly (and visa versa), the test is said to have high criterion-related validity. Content-related validation is a demonstration that the content of the test reflects important job-related behaviors and measures important job-related knowledge or skills.

Construct-related validity is evidence that a test measures the constructs or abstract characteristics that are important to successful performance of the job. For psychological tests used in selection, a test's criterion-related validity is usually the variable of interest to researchers, and it is the validity coefficient—the actual correlation coefficient between a test score and some job performance criterion—that is referred to when validity is discussed in HR literature.

Having evidence of the validity of selection tests is essential for any organization using such tools. Collecting these data is the principal way companies demonstrate that they have met the Uniform Guidelines' requirements should hiring procedures result in adverse impact (i. e. disproportionate hiring outcomes) against protected groups. Many experts and personnel selection specialists believe that test validity can be

attenuated or even sacrificed to reduce adverse impact. Often, a practitioner is faced with a choice among tests having very different costs, degrees of validity, and fairness.

The Uniform Guidelines provide guidance on making such choices: When two procedures are available that are valid and reliable and that serve the company's interest in efficient and trustworthy workmanship, the company should use the procedure that has been demonstrated to have the " lesser adverse impact. The selection professional is faced with the obligation of developing a test that has high validity and minimal or no adverse impact in an environment where racial, gender, and social groups may have different ability distributions and in which poor selection outcomes would adversely impact the prospects of the hiring organization. Consequently, the procedure known as the Golden Rule, which consists of selecting items on the basis of reduced adverse impact and compromising construct validity (and reliability), has become widely held. There are many different perspectives from which to view fairness in personnel selection, and each has implications for psychological testing and decision-making. Dreher and Sackett identified five models of fairness, including pure quota and culture-free models in which fairness in both is defined based on how well selection ratios represent protected group proportions in a relevant labor market.

Other definitions, such as the models developed by Cleary and by Thorndike, incorporate the issues of single-group validity versus differential validity and differential prediction. In these models, the fairness of a selection test is determined by examining the regression line describing job performance as a function of test scores, which is the line of best fit through a scatter-plot of

<https://assignbuster.com/employee-selection/>

data points that show an individual's test scores on the horizontal axis and job performance measures on the vertical axis. Typically, a test will only be considered to be fair when the regression lines for minority and non-minority groups are the same. When regression lines are not the same for minority and non-minority groups, then using a test in the same way for both would not be considered fair. Psychological tests and inventories have been the most frequently used predictors in I/O psychology” (Muchinsky, 2006 p.

99) One potential problem for the use of personality inventories in personnel selection is the frame-of-reference problem. Most available personality tests are composed of items referring to behavioral tendencies, attitudes, relationships, preferences, and social skills. Thus, these items represent individuals' characteristic adaptations that in aggregation are commonly thought to generalize across situations. It is assumed that individuals respond to items with an indication of their propensity to behave, feel, relate, and so on, in a general way across situations (or at least they present an image of how they wish to be regarded across situations). This may not be the case, however, for all job applicants completing personality inventories, as some may adopt a specific frame of reference in answering items.

Some job applicants may feel that the employer only wants to know (or only has the right to know) how they are likely to behave, feel, relate, and so on, at work. For example, general context items reflecting the Agreeableness factor (i. . , a trait from the Big Five model of personality) may be answered by some job applicants in relation to their self-perceived agreeableness at

work; specifically and by others in a way that reflects their general propensity for agreeableness across all situations.

With the evolution and development of appraisal systems, a number of tools and techniques of performance appraisal have been developed. Firstly, there are graphic-rating scales, which compare individual performance to an absolute standard. In this method, judgments about performance are recorded on a scale. This is the oldest and most widely used technique. This method is also known as linear rating scale.

Rating scales are of two types: continuous rating scales and discontinuous rating scales. Another popular and effective method of performance appraisal is employee ranking. Under this method the employees are ranked from best to worst on some characteristics. The rater first finds the employee with the highest performance and the employee with the lowest performance in that particular job category and rates the former as the best and the latter as poorest.

Then the rater selects the next highest and next lowest and so on until he rates all the employees in that group. 60 degree feedback is the most comprehensive and costly type of appraisal. It includes self ratings, peer review, and upward assessments; feedback is sought from everyone. It gives people a chance to know how they are seen by others; to see their skills and style; and may improve communications between people. 360 degree feedback helps by bringing out every aspect of an employee's life.

Cooperation with people outside their department, helpfulness towards customers and vendors, etc. may not be rewarded by other types of

appraisal. This system also helps those who have conflicts with their manager. 60 degree feedback generally has high employee involvement and credibility; may have the strongest impact on behavior and performance; and may greatly increase communication and shared goals. It provides people with a good all-around perspective.

The value of self-assessment lies in its ability to make the rater take responsibility for their own performance and development. When an individual or group participates in self-evaluation to create a development plan, there is an increased level of commitment to the goals formulated as a result of the assessment (Atwater, 1998; Wilson & Pearson, 1995). When used as part of the performance appraisal process, self-evaluation also enhances the rater's dignity and self-respect. The rater is an active participant in evaluating their performance and is not at the mercy of a supervisor or other evaluator. Self-assessment works by helping the rater internalize the need for change and performance improvement.

Individuals are encouraged to take responsibility and ownership for their own improvement; the motivation for change comes from within rather than being imposed from outside. In peer assessment, members of a group appraise the work of others in the group. This is done by peer nomination, which is nominating a specified number of group members as being highest on a particular dimension of performance. Another option is peer ratings in which each group member rates the others using rating scales.

Finally, peer ranking is a technique where each member ranks all others from best to worst on one or more performance dimensions. Issues arise in

this kind of assessment due to friendships or fear of retaliation. Interviews are the most frequently used employee selection process in modern business. Research also shows that unstructured interviews are almost completely indefensible on scientific or legal grounds. Structured interviews, which use a standard script tailored to the demands of the job in question and a standard way to interpret answers, are in fact defensible on scientific and legal grounds. The problem with structured interviews concerns “interview drift”-interviewers get bored with the standard procedure and drift back into unstructured interviews and all the problems that they entail.

Integrity tests are typically used to screen entry-level employees for honesty, dependability, and willingness to follow rules. These tests have two attractive features. First, they are scientifically defensible – they predict job performance. And second, they are legally defensible – they don’t discriminate against minority job applicants. The problem concerns their narrow focus – they only evaluate a person’s willingness to follow rules.

They have nothing to say about potential for customer service work, for working as part of a team, for exercising leadership or taking initiative, or for thinking creatively. BARS refers to Behaviorally Anchored Rating Scales. It was developed by Smith and Kendall to provide a better method of rating employees. It differs from standard rating scales in one central respect, in that it focuses on behaviors that are determined to be important for completing a job task or doing the job properly, rather than looking at more general employee characteristics (e. g. personality, vague work habits).



So, rather than having a rating item that says: Answers phone promptly and courteously, a BARS approach may break down that task into behaviors. Human performance is largely a qualitative phenomenon and cannot be precisely measured quantitatively. Though quantifying performance elements can make the overall process more objective, but not exact or perfect. After all, human behavior is a complex phenomenon – far more perplexing than the measurement problems in Physics – weight and length etc. High precision in performance measurement for most of the jobs is a mere illusion. Apart from the inexactness of measurement, a few cognitive and perceptive problems often cause the raters to make significant errors in judgment.

For instance, a halo effect occurs when a rater attaches too much significance to a single factor of performance and gives similar ratings on other performance elements. Thus overall evaluation is significantly influenced by a single factor. Such a perception can undermine the importance of other elements and leads to an unbalanced performance assessment of the individual. People differ also in their tendency to evaluate people or performance. Some supervisors are very strict or conservative in their ratings and generally give low scores in their evaluations.

This tendency may make high performers attain somewhat average ranking and average performers appear as poor performers. Raters with such tendency are known to have a strictness bias. Performance appraisal involves assessment of employee performance for a specific period – quarterly, annually etc. People may not perform uniformly throughout that period. We all face highs and lows and demonstrate variance in performance

due to numerous factors. It is therefore very important to review performance demonstrated throughout the period under consideration.

Often however, recent events tend to overshadow the overall performance. People do have “ short memories”. Thus a person who has worked very hard and excelled throughout the year, but for some inadvertent reasons had faced performance issues in the last weeks or month may at times get a poor appraisal from their supervisor, showing a recency bias. Personal beliefs, attitudes, assumptions, experiences, preferences and lack of understanding about a person, class or a phenomenon can lead to an unfair evaluation which is off from reality.

We all suffer from these shortcomings, consciously or unconsciously, while making everyday judgments about people, things, events etc. It is especially important to be aware and sensitive to possible biases, prejudices and stereotypes while making judgments about employee performance. While many of the prejudices operate covertly and unconsciously, others strike us through conscious thoughts and feelings. Understanding common biases and being on guard while appraising can significantly raise the objectivity of the evaluation process.

References Barrett, G. V. , Alexander, R. A. , & Doverspike, D. (1992).

The implications for personnel selection of apparent declines in predictive validity over time: A critique of Hulin, Henry, and Noon. *Personnel Psychology*, 45, 601. David I. George, Mike C. Smith (1990); *Public Personnel Management*, Vol.

19, Schmit, Mark J. (1995), Frame-of-reference effects on personality scale scores and criterion-related validity. B. , Journal of Applied Psychology, 0021-9010, Vol. 80, Issue 5 Hunt, E. B.

(1995). Will we be smart enough: A cognitive analysis of the coming workforce.. New York: Russell Sage Foundation. Muchinsky, P. (2006).

Psychology Applied to Work. (8th ed. ). (pp. 99).

California: Thomas Wadsworth Ones, D. S. , Viswesvaran, C. , & Schmidt, F. L.

(1993). Comprehensive meta-analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance [Monograph]. Journal of Applied Psychology, 78, 679–703.