

SELECTION

Except in small organizations, the human resource department assumes the major responsibility for employee selection. The human resource department normally reduces the field of applicants to three or four possible candidates. Line managers then interview these candidates and make their selection.

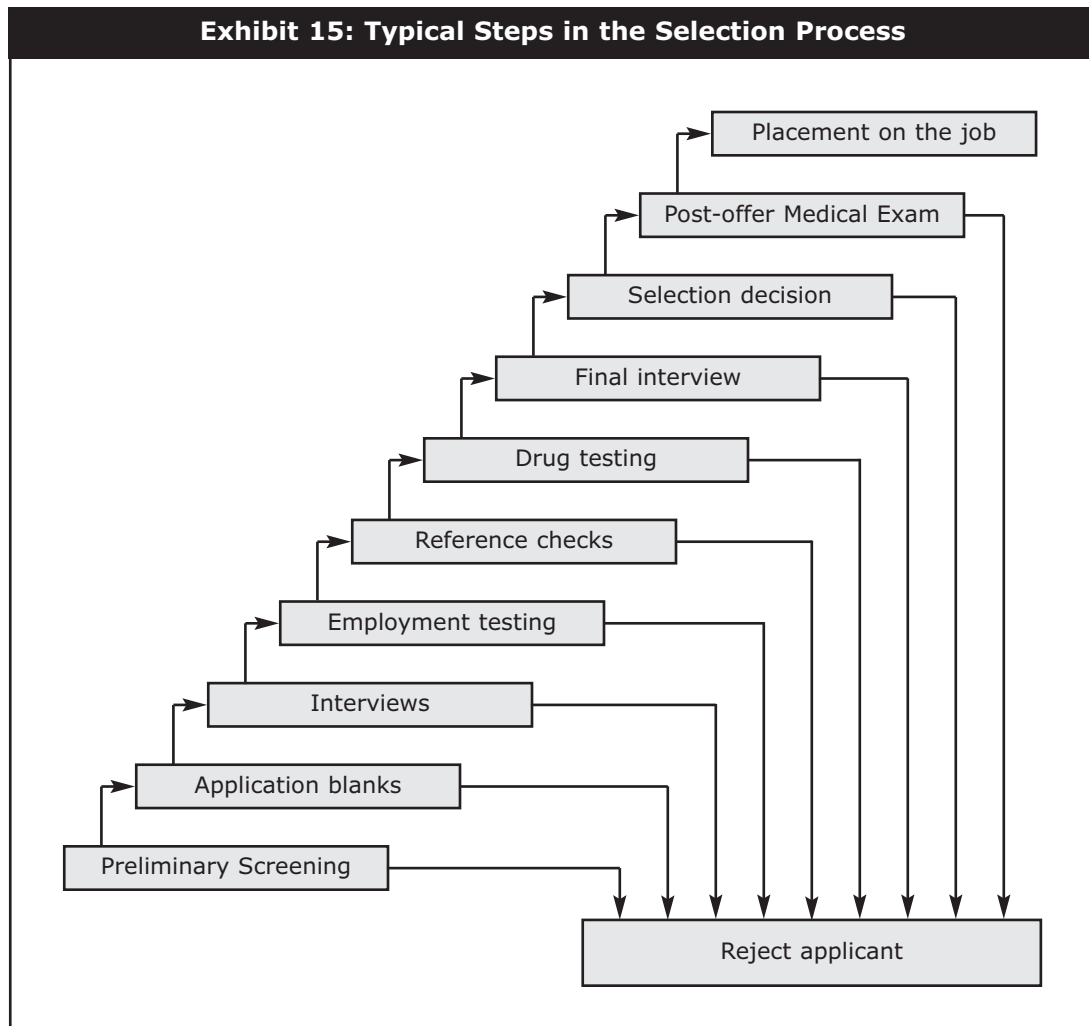
Application Process

The process of making an informed hiring decision depends largely on two basic principles of selection.

1. Past behavior is the best predictor of future behavior. Knowing what an individual has done in the past is the best indication of what the individual is likely to do in the future. This principle is not deterministic; knowing what people have done in their past is not an absolutely accurate indication of what they will do in the future. Nevertheless, in making selection decisions it is a good rule to follow.
2. Organizations should collect as much *reliable and valid data* as is economically feasible and then use it to select the best applicants. Reliable data refers to information that is repeatable and consistent. Valid data refers to information that indicates how well employees will perform their jobs. In some situations very little information needs to be collected regardless of its reliability and validity. Extensive information is unnecessary when a job is simple and can be performed by almost any applicant or when the cost of making a bad hiring decision is negligible. However, as jobs become increasingly difficult to staff with competent employees and as the costs of making a poor hiring decision increase, the collection of reliable and valid information becomes very important.

The selection process is a sequential procedure involving some or all of the steps illustrated in Exhibit 15. Each step in the process is typically treated as a hurdle that systematically screens the number of employees advancing to the next step. To survive the process and be placed in the job, an applicant must successfully pass each hurdle. Most organizations reject undesirable applicants at each step of the process to reduce the burden of keeping track of a large number of applicants, but some organizations have all applicants go through the entire selection process, waiting until the end to choose the best candidate. Each step in the selection process should be designed to obtain specific, useful information for making a hiring decision.

The order of the steps in the application process should be determined by a cost-benefit analysis; the most costly and subjective steps should be placed at the end of the process when fewer applicants remain in the pool.



Reliability

All selection information should be reliable. **Reliability** refers to consistency of measurement or repeatability. A selection instrument, such as an achievement test, is said to be reliable if individuals obtain essentially the same scores each time they take the test. The higher a measuring instrument's reliability, the greater the confidence that can be placed in the instrument.

The Importance of Reliability

The issue of reliability has been considered extensively with respect to personnel testing. Test publishers include evidence of reliability in the examiner's manuals and in advertisements about the tests. However, reliability concerns more than just personnel testing; the reliability of every selection variable should be examined, including application blanks, interviews, reference checks, and physical exams. The issue of reliability is also important for other human resource functions, especially performance evaluations. If a supervisor cannot reliably evaluate the performance of a subordinate, the evaluation should not be used for human resource decisions such as firing, promoting, or granting pay increases.

Reliability is an important issue in employee selection because it serves as a limiting constraint on validity. If a measuring instrument is not reliable, it cannot be valid; unreliable information cannot be used to predict performance on the job.⁹⁴ Unreliable data are essentially capricious random numbers that do not really measure anything. Reliability is typically indicated by a correlation coefficient, and correlations below .70 are generally considered unacceptable. Some of the major reasons why selection information may be unreliable include:

1. The selection instrument may be ambiguous and unclear. For example, items on an application blank may not be specific, or questions in an interview may not be clear. Similarly, test questions that are ambiguous may lead to random responses that do not measure a consistent, repeatable personality characteristic.
2. The person using the measuring instrument may not have a clear perception of the behavior being measured or a well-defined standard to use as a basis for making an evaluation.
3. The behavior being evaluated may be an unstable phenomenon that changes from time to time, such as personal feelings, rather than a stable personality characteristic.

Reliability is not the same as accuracy. The measuring instrument may be reliable and still not be accurate. For example, a cloth measuring tape may stretch over time and not provide an accurate measurement of length. However, the tape would still be a very reliable measuring instrument because it produces the same measurements time after time.

Operational Definitions

The reliability of a measuring instrument is estimated by different types of tests. These testing procedures are referred to as *operational definitions*, since the reliability is defined by specifying the operation used to estimate the reliability. The four operational definitions of reliability include (1) test-retest reliability, (2) alternate-forms reliability, (3) split-halves reliability, and (4) inter-rater reliability. Each of these methods produces a correlation coefficient, which is a number that can range between plus one (+1.00) and minus one (-1.00).

Test-retest Reliability: One of the most obvious ways to test the reliability of a measuring instrument is to use the same measure twice on the same sample of people and determine if the second measures are similar to the first. This procedure is called **test-retest reliability**, or *coefficient of stability*.⁹⁵ If the measuring instrument is reliable, the individual's score on the second measurement should be essentially the same as that obtained on the first measurement. The test-retest reliability of most selection instruments can be examined quite easily. For example, the reliability of employment tests and application blanks can be checked by administering the instruments when the applicants first apply and then asking the applicants to complete the same instruments again one week later. If the responses are essentially the same, the instruments would be considered reliable. The reliability of interviews and reference checks also can be examined by repeating them to determine if the same evaluations were made both times.

When the test-retest reliability of an instrument is examined, an important consideration is how much time should elapse between the first and second measurement. A useful rule is

that enough time should pass so that the individuals do not remember the responses they made the first time, but not so much time that the characteristic being measured has an opportunity to change. A one-week interval is probably sufficient for testing the reliability of application blanks, interviews, and some personnel tests. However, a two- or three-month interval may be necessary for testing the reliability of a mental-ability test.

Alternate-forms Reliability: The reliability of a measuring instrument can be estimated by developing an alternate form of the instrument and then correlating the responses of a sample of people to both forms. Many test publishers, for example, develop two or more forms of the same test. Both Form A and Form B of the same test can be given to a group of applicants, and the scores can be compared to see if the scores for each individual are essentially the same. Although the tests are not identical, the number of questions and the

Exhibit 16: Estimating the Reliability of a Mental Ability Test Using the Alternate-Forms Method (X=Form A and Y=Form B) N=25 Students

Students	X	Y	X ²	Y ²	XY
1	18	17	324	289	306
2	19	17	361	289	323
3	30	28	900	784	840
4	26	25	676	625	650
5	19	22	361	484	418
6	23	22	529	484	506
7	28	25	784	625	700
8	24	27	576	729	648
9	19	17	361	289	323
10	18	20	324	400	360
11	22	23	484	529	506
12	31	31	961	961	961
13	26	25	676	625	650
14	24	26	576	676	624
15	17	18	289	324	306
16	19	21	361	441	399
17	20	17	400	289	340
18	23	23	529	529	529
19	28	29	784	841	812
20	30	31	900	961	930
21	32	33	1,024	1,089	1,056
22	21	22	441	484	462
23	20	18	400	324	360
24	25	23	625	529	575
25	26	25	676	625	650
Totals (Σ) =	588	585	14,322	14,225	14,234
$r_{xy} = \frac{N\sum(XY) - (\sum X)(\sum Y)}{\sqrt{N\sum X^2 - (\sum X)^2} \sqrt{N\sum Y^2 - (\sum Y)^2}}$ $= \frac{25(14,234) - (588)(585)}{\sqrt{(25)(14,322) - 588^2} \sqrt{(25)(14,225) - 585^2}} = 0.92$					

nature of the items are essentially equivalent. If the scores are highly correlated, the test is said to be a reliable instrument. **Alternate-forms reliability**, which is sometimes called the *coefficient of equivalence*, avoids the problem of having to decide how long to wait to administer the second form.⁹⁶ Both Form A and Form B can be administered at the same testing session.

An illustration of alternate forms reliability is presented in Exhibit 16, which shows the scores on Form A and Form B of a mental-ability test for 25 individuals. The test publisher advertises that Form A and Form B are equivalent measures of mental ability and the high correlation coefficient of .92 supports this claim.

Split-halves Reliability: Another method of estimating the reliability of a measuring instrument is to split the instrument in half, forming two separate scores, and then to correlate the scores for each individual to determine if comparable scores have been obtained. This **split-halves** reliability method is similar to alternate-forms reliability except that one test is divided into two parts rather than developing two alternate tests. The split-half method, which is sometimes called the *coefficient of internal consistency*, can only be used if the test is measuring a single dimension.⁹⁷ The two halves must theoretically be measuring the same personal characteristic to be meaningful.

Inter-rater Reliability: Two people observing the same behavior may or may not evaluate it the same. The degree of consistency between the scores assigned by two different observers is referred to as **inter-rater reliability**, or **conspect reliability**. Inter-rater reliability is especially important in assessing the reliability of interviews and performance evaluations. If two interviewers talk with a group of individuals and independently evaluate them, their evaluations are said to have high inter-rater reliability if they agree. Unfortunately, the inter-rater reliabilities of interviews and performance evaluations are often low.⁹⁸

Validity

Validity refers to the extent to which a *predictor* variable is correlated with a *criterion* variable. (In experimental research, validity has a slightly different meaning, as is discussed in Unit 3.) *Predictor* variables refer to all the kinds of information that are collected as part of the employment process, such as application blank data, interview data, test scores, and personal references. A *criterion* variable refers to a measure of job performance, such as a measure of productivity, absenteeism, tardiness, supervisory evaluations, or any other information that indicates the degree of success on the job.

Operational Definitions

The five methods for evaluating the validity of a selection procedure are (1) predictive validity, (2) concurrent validity, (3) content validity, (4) construct validity, and (5) synthetic validity.

Predictive Validity: The procedure used for determining **predictive validity** consists of gathering the predictor information on job applicants at the time they apply for employment without using this information in the hiring decision. After all new employees have been oriented and trained, the criteria data (performance measures) are obtained, and a correlation coefficient is then computed between the predictor data and the criteria data. If the validity

coefficient is significantly greater than zero, the predictor is considered valid and it can now be used to select the best employees.

This procedure also is called the “follow-up” method of validation since subsequent performance is evaluated and used to assess the soundness of the selection procedure. Predictive validity is the most valuable kind of validity information since the method used to collect the predictor data is identical to the way the data will be collected in the future. Furthermore, evidence suggests that predictive validity persists over time regardless of the effects of training and experience, at least with ability tests. Individuals who are successfully predicted to be good performers immediately, continue to be good performers several years later.⁹⁹

Concurrent Validity: Sometimes employers do not want to wait until after orientation and training periods to validate their selection procedures. Moreover, unless employers are hiring many new employees, they may have to wait a long time to obtain a sample size large enough to conduct a predictive-validity study. In these situations an alternative procedure is to conduct a **concurrent-validity** study, which is sometimes called the “present-employee” method. This method involves collecting the predictor data and the criteria data on a sample of present employees and then correlating them to determine if a relationship exists. The difference between predictive and concurrent validity is illustrated in Exhibit 17.

