Chapter 05
Psychological Testing

S.P MANDALI'S weschool Welingkar Education

Learning Objectives

After completing this chapter you will be able to:

- Know the fundamental principles and assumptions of psychological testing.
- Define methods used to develop, evaluate and utilize psychological tests.
- Learn the application of psychological testing in organizations.
- Describe the ways psychological tests can be similar to and different from one another, their advantages and their limitations.
- Understand the terminology in the field of psychological testing .

Structure

1 Introduction to psychological testing

1.1 The purpose of psychological testing.

1.2 The Advantages and Disadvantages of psychological testing

1.3 The factors influencing psychological testing

2 Test Measurement

3 Test Development

4 Test reliability

4.1 Test-Retest

4.2 Alternate Forms

4.3 Internal Consistency

4.4 Inter-Rater

Structure

5 Test Validity

5.1 Criterion Related Validity

5.1.1 Concurrent Validity

5.1.2 Predictive Validity

5.2 Content Related Validity

5.3 Construct Related Validity

5.4 Face Validity

Structure

6 The Types of Psychological Testing

6.1 Intelligence Test

6.1.1 The Stanford-Binet Intelligence Scale

6.1.2 The Wescher Adult Intelligent Scales

6.1.3 The Otis Test

6.1.4 The Chicago Test

6.2 Ability Test

6.2.1 The Minnesota Spatial Relations Test

6.2.2 The Bennett Mechanical Comprehension Test

6.2.3 The Detroit Clerical Test

6.2.4 The Reaction Time Test

6.2.5 The Manual Dexterity Test

Structure

6.3 Aptitude Test

6.3.1 The Differential Aptitude Test (DAT)

6.3.2 The General Aptitude Test Battery – GATB

6.3.3 The SRA Mechanical Aptitudes Test

6.4 Personality Tests

6.4.1 The California Psychological Inventory – CPI

6.4.2 The Minnesota Multiphasic Personality Inventory – (MMPI)

6.4.3 The Personality Adjective Checklist – PACL

6.4.4 The Cross Cultural Adaptability Inventory - CCAI

6.4.5 The NEO –PI-R (Neuroticism, Extraversion and Openness Personality Inventory

Structure

1. Introduction to Psychological Testing

Even though everyone faces different examinations in daily life, when it comes to psychological tests, there is (unwarranted) fear in the minds of candidates.

Psychological testing practice, though, is as old as psychology itself. The term 'mental test' was introduced by Cattell back in 1890. During world wars the intelligence tests were administered while recruiting soldiers.

In 1960s the psychological tests were opposed as tests had very little relationship to the job for which they were administered; and too many personal details were checked that invaded privacy of candidates.

1. Introduction to Psychological Testing

The restraint on tests lasted well into 1970s and until more evidence of validity for tests had become available. By mid 1980s, testing was back in full swing and both intelligence and personality testing began to appear with greater frequency.

Robert Guion (1998) defined a test as 'an objective and standardized procedure for measuring a psychological construct using a sample of behavior.'

Tests cover paper and pencil tests, interviews, a work sample test and even an application blank.

1. Introduction to Psychological Testing

Psychological Test

An objective and standardized series of questions for measuring person's knowledge, skills, abilities or other characteristics(KSAO).

Psychological testing is also called psychometrics.

The term 'objective' implies quantification. At the end of the test candidates receive a score which can be just pas / fail or a score on some range. The simple process of assigning a score is quite different from interpreting it.

1. Introduction to Psychological Testing

The definition is broad enough to cover many different types of content including

- **cognitive ability**
- **personality**
- **values**
- **communication skills**
- **interpersonal skills and**
- **technical knowledge.**

Most employers look at several attributes (KSAOs) using several techniques.

1. Introduction to Psychological Testing

A score of 88% signals excellent performance in the A range, if only two three candidates in the group scored above 85%, but not so if others, except you, scored over 94%. Interpreter can compare your score to that of previous groups who undertook the same test.

The development of test norms is very technical but it is simply important to be aware that while a test produces a 'score' there is a need to interpret or give meaning to that score. The task of interpretation can be entrusted only to a person formally trained in the area of psychological assessment.

1. Introduction to Psychological Testing

Psychological testing, if it is to be done ethically and effectively, is not a simple process. Psychological tests are used to make important decisions by both the individuals and the organizations.

Today the content and the process of employment testing is varied and encouraging. These tests vary according to

- **their mode of administration.**
- **their level of standardization.**
- **structure.**
- **their costs.**
- **their administrative ease, etc.**

1. Introduction to Psychological Testing

Industrial Psychologists have identified many different attributes that appear to contribute to work performance and also many different methods for assessing them. One of the methods that the industrial psychologists use is the Test Battery.

A test battery is a collection of tests rather than a single test. The tests in the battery are generally of different attributes. These attributes may be within a single area, such as cognitive battery including subtests of reasoning, memory and comprehension.

1. Introduction to Psychological Testing

Or these attributes may be from conceptually different areas, such as a battery that includes a measure of cognitive ability, a personality test, a physical ability test and a test of vocational interests.

The term battery usually implies that all the tests will be taken either in a single testing period or over a very short period of time.

TEST BATTERY

Collection of tests that usually assess a variety of different attributes.

1. Introduction to Psychological Testing

Assumptions of Psychological Tests

When using psychological tests, we must make important assumptions, including

i. **A test measures what it says it measures**
ii. **An individual's behavior and, therefore, test scores will remain stable over time.**
iii. **Individuals understand test items similarly**
iv. **Individuals can report accurately about themselves**
v. **Individuals will report their thoughts and feelings honestly.**
vi. **Test score is equal to ability plus some error.**

1. Introduction to Psychological Testing

Ethical Standards for Psychological Testing

Because testing is important but often misunderstood most professional societies have a set of professional practice guidelines known as ethical standards. These codes affirm the importance of respect for individuals, establish the need to safeguard individual dignity and privacy and condemn unfair discriminatory practices.

Focus here is on

- **maintaining confidentiality and anonymity**
- **obtaining informed consent**
- **using test appropriately.**

1. Introduction to Psychological Testing

Ethical Standards for Psychological Testing

- Ethics reflect what is considered “right” or “wrong” of a society or culture.
- Most professional bodies have a set of ethical standards that are adopted by their members.
- The American Psychological Association(APA) and others have ethical standards that protect the rights of individuals who are the recipients of psychological testing.
- Most important, test takers have the right to privacy, anonymity and informed consent.

1. Introduction to Psychological Testing

Appropriate Use of Psychological Tests

- Test publishers should follow professional standards and guidelines by marketing tests truthfully, selling tests only to qualified users, providing evidence of validity, developing a comprehensive test manual for each test and maintaining test security.

- Test users have the responsibility to ensure that everyone who uses the test has the necessary training and experience to carry out his/her obligations in regard to the test.

1. Introduction to Psychological Testing

Appropriate Use of Psychological Tests

- A test user is anyone who participates in purchasing, administering, interpreting or using the results of a psychological test
- A test taker is a person who responds to test questions or whose behavior is being measured.

The social and legal implications of psychological testing: some individuals are concerned that these tests are biased & do not result in the correct institutional decisions. This controversy is centered around aptitude, intelligence and integrity tests & has influenced social movements and laws on their use.

1. Introduction to Psychological Testing

Psychological Assessments, Tests, Measurement & Surveys

- Psychological assessment involves multiple methods for gathering data.
- Measurement is the assignment of numbers according to rules.
- A test is a measurement when the behavior sample can be expressed as a numerical score.
- Surveys focus on group outcomes & tests focus on individual differences.
- Survey results are typically reported at the question level and test results are typically reported as an overall score.

1. Introduction to Psychological Testing

Psychological Assessments, Tests, Measurement & Surveys

- **Tests and Testing**
- **Speed test:** has a rigid and demanding time limit so most takers will be unable to finish the test in allotted time
- **Power test** : has no rigid time limit. Enough time is provided for a majority of the takers to complete all the test items.
- **Group test**: can be administered to large groups of individuals. Therefore it is valuable in reducing the costs of testing, both in time and money.
- **Individual test:** test given on an individual basis

1. Introduction to Psychological Testing

Psychological Assessments, Tests, Measurement & Surveys

- **Tests and Testing**

- **Paper and pencil test : one of the most common forms of testing that requires no manipulation of any objects other than the instrument used to respond**
- **Fairness: value judgment about actions or decisions based on test scores.**
- **Bias: technical and statistical term that deals exclusively with a situation where a given test results in errors of prediction for a subgroup.**
- **Assessment: appraising or estimating the magnitude of one or more attributes in a person.**

1.1 The purpose of psychological testing

Psychological testing is used mainly at two stages –

1. **first in the pre-employment phase;**
2. **and second in various performance appraisal tools (after employment).**

Formal testing begins with an interview wherein information regarding suitability of the candidate is assessed by the interviewer through a standardized format of questions in a set order.

1.1 The purpose of psychological testing

Psychological testing is used in occupational context in many ways including

- **Employee selection.**
- **Career development.**
- **Leadership development.**
- **Talent identification.**
- **Performance management**
- **Outplacement.**
- **Assessment of job-person fit.**
- **Team development.**

- to name a few

1.1 The purpose of psychological testing

In more recent years organizations have benefited from speed, efficiency and cost effectiveness of internet testing to support implementation of systematic and structured selection and development processes at individual and group levels.

Structured selection processes including those supported by online psychological tests with known reliability and validity are considered to be fairer to job candidates than alternative interview and resume based selection methods.

With internet testing, psychological testing and tests have become popular and more accessible.

1.1 The purpose of psychological testing

Organization benefits

1. Underlines personality strength / weaknesses of the candidate relative to the demands of the role and competencies required.
2. Assists in the process of differentiating between equally favorable candidates for a position.
3. Provides objective group ranking options and shortlists for large intakes.
4. Supports teambuilding, leadership development, career development and work-life balance initiatives.

1.1 The purpose of psychological testing

Candidate benefits

1. Provides insights into their strength and weaknesses and allows them to look at upgrading skills.
2. Assists them assessing whether the role they are applying for is a good fit in terms of their personality and intellectual abilities.
3. Serves as a guideline for career development and career options in the future.
4. Provides them with a strategy of how to manage their performance in an organization.
5. Gives them a measure of their relative standing to the benchmarks used.

1.2 Psychological Testing
the Advantages and Disadvantages

Using psychological and personality tests for job recruiting can have several advantages. Screening prospective employees can reduce the amount of time that hiring managers spend conducting interviews.

Administering tests before a candidate is selected for a particular job opening helps to reduce business costs and employee turnover by allowing employers to hire candidates who are ideally suited for a job specification.

Employment tests are easy to administer.

Employment tests can be delivered online or through telephones.

1.2 Psychological Testing
the Advantages and Disadvantages

Results you get are objective rather than subjective.

Answers are either right or wrong , results cannot be faked and tests create a level playing field for comparing candidates.

Verbal Reasoning and Numerical Tests have shown high validity for a wide range of jobs. Validity rises with increasing complexity of job. Combinations of aptitudes tests have higher validities than individual tests alone. Can be administered in group settings and scoring may be completed using computer scanning.

1.2 Psychological Testing
the Advantages and Disadvantages

Yet there are some problems.

Job applicants can easily manipulate employee personality and psychological responses to create positive outcomes.

Non- minorities score one standard deviation above minorities.

Gender differences result in lower score for female candidates.

All abilities are not tested.

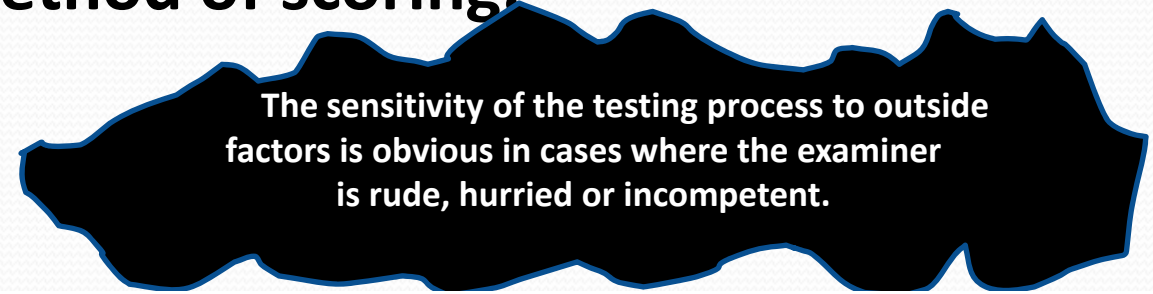
Like all exams, after all, it tests a one off experience only.

1.3 The Factors Influencing Psychological Testing

Psychological Testing is a dynamic process influenced by many factors.

Although raters strive to ensure that test results precisely reflect the traits or capacities being assessed, many factors can sway the outcome of Psychological testing such as

- **the manner of administration,**
- **the characteristics of the tester,**
- **the context of testing,**
- **the motivation and experience of the examinee and**
- **the method of scoring.**



1.3 The Factors Influencing Psychological Testing

Organizational factors.

Administrator Qualifications **– many psychological tests are difficult administer. Performing the tests requires knowledge about how statistical concepts such as a t-score relate to the validity of test results.**

Test Material Secrecy **– Some psychological tests are only effective because their methods are secret. Employees who have a prior knowledge about the questions on an exam such as a psychological survey, an employee can provide answers which employer wishes to receive.**

1.3 The Factors Influencing Psychological Testing

Organizational factors.

Test Expense **– many psychological tests are under copyright and may not be used without the permission from a testing center or the psychologist who created the test. Test providers may charge a fee to administer the test. Using test without permission / payment may violate law.**

Test Environment– **a supervisor must control the environment in the test center. Performing psychological tests over phone or on employee's computer does not provide desired control.**

1.3 The Factors Influencing Psychological Testing

Applicant's Mental or Physical State

Level of Motivation.

Alertness

Anxiety at the time of testing.

ALL HAVE INFLUENCE ON THE RESULTS.

1.3 The Factors Influencing Psychological Testing

Assessment factors

Instructions to Applicants

Time Limits

Use of calculators or other resources.

ALSO HAVE INFLUENCE ON THE RESULTS.

Some more factors

1.3 The Factors Influencing Psychological Testing

Measurement Conditions

Lighting
Temperature
Noise Level
Visual Distractions

Scoring Procedure

Raters who evaluate applicant performance in interview
Assessment center Exercises

Writing Tests

2. Test Measurement

Measurement is the keystone of both the science and practice of industrial psychology. Measurement is essential because it is the means to describe, predict, explain, diagnose and make decisions about various tasks in the organization.

Without standard measurement, the practice would be haphazard.

Measurement is the assignment of numbers to attributes or properties of people, objects, or events based on a set of rules (Stevens, 1968)

2. Test Measurement

First measurement focuses on attributes of people, objects or events and not actual people, objects or events.

Second measurement uses a set of standardized, clear and easy to apply rules to quantify these attributes.

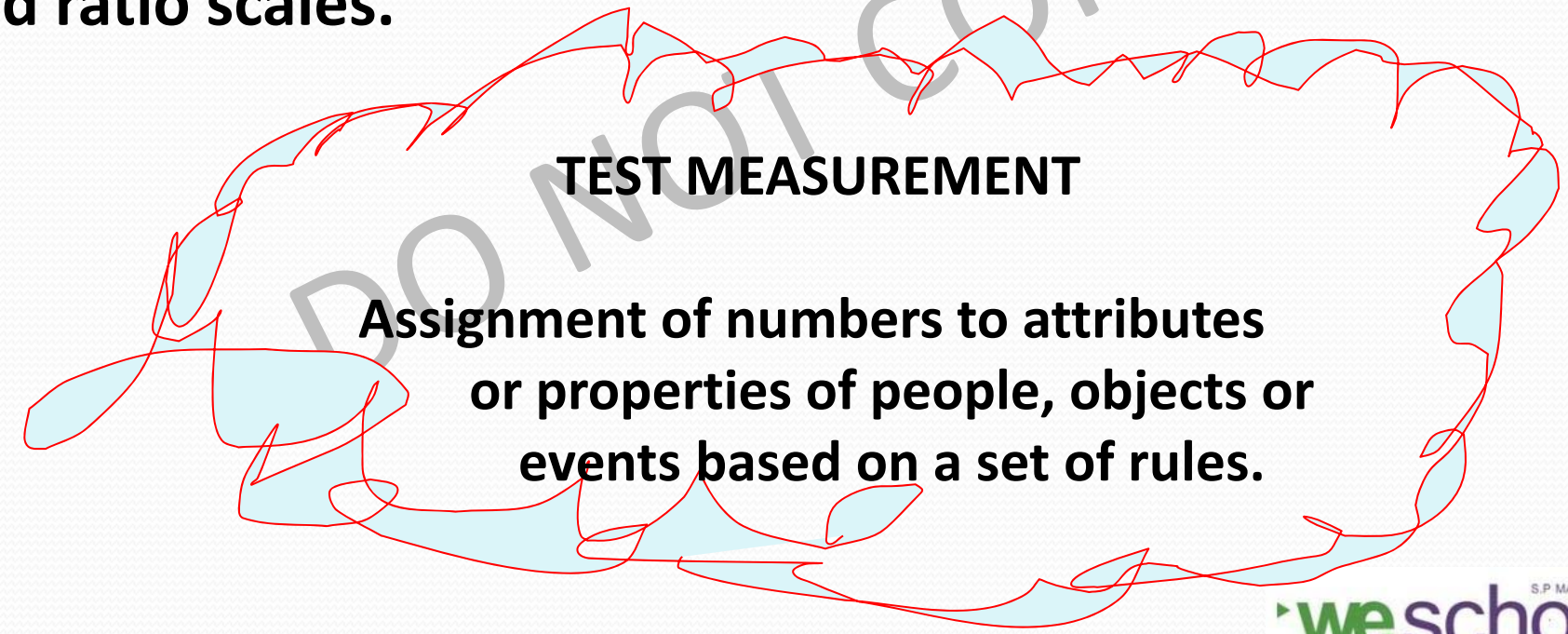
Third measurement consists of two components, scaling and classification.

Scaling is the assignment of numbers and classification refers to defining if people, objects or events fall into the same or different category based on given attribute.

2. Test Measurement

A very important series of articles by Stevens introduced the idea that there are four categories of scaling or levels of measurement that are used in psychology.

These are Nominal scales; ordinal scales; interval scales and ratio scales.



2. Test Measurement

Nominal scales

The nominal level of measurement is often also known as categorical data. It is the most basic and it involves assigning numbers to the data such as labels (telephone numbers) or categories of objects. Keeping record of sex of participants in a workshop will require to put the in category 01 – females and category 02 –males.

Here numbers 01 or 02 have no meaning other than labels. One is not greater (or smaller) than other, they are just different. Calculating their mean or arithmetical average will be futile.

Ordinal data records categories

2. Test Measurement

Ordinal data record information about rank order of scores.

Ordinal Scales

If we arrange objects or individuals into a rank order, we are using an ordinal scale.

Data sorted in this way are sometimes also referred to as ordered categorical data. Five best TV serials selected by you is an example of ordinal scale. The serial you like most will have rank 01.

Another example would be the first, second and third runner to complete the marathon. Second runner may take ten more seconds to cross the finishing line while the third just two extra seconds. Thus ordinal scales do not have equal intervals.

2. Test Measurement

Interval scales

Interval data represents a much higher level of measurement than ordinal data. When data are measured on an interval scale, we know about the rank order (as with ordinal data) but we also know about the intervals – that is the differences.

When we measured data on an ordinal scale we did not know anything about the intervals – that is the differences. We did not know time elapsed between the first and second runner. We did not know the size of the interval.

2. Test Measurement

Interval scales record the order of data points and the size of the intervals in between data points.

Interval scales

The centigrade scale of temperature measurement is an example of an interval scale. The intervals between adjacent values in the centigrade scale are equal; it is reasonable to add and subtract the numbers on the centigrade scale.

Temperature change from $10℃$ to $20℃$ is exactly the same size as the change from $40℃$ to $50℃$.

However, interval scales do not have an absolute zero point. For example, $0℃$ does not mean that the thermometer is measuring no heat at all, the 0 mark is simply a point on the scale

2. Test Measurement

Ratio scales

The ratio scale has all the properties of the interval scale – equal intervals that can be added and subtracted. The important difference is that the ratio scale, unlike the interval scale, does have a true zero point.

The implication of having a true zero point is that numbers on the scale, as well as being added and subtracted, can also be multiplied and divided – we can say that one measurement was twice as big as another.

This cannot be done with the scale that does not have 0.

$20℃$ is not twice as hot as $10℃$!

2. Test Measurement

A ratio scale is an interval scale, with a true zero point.

Ratio scales

As an example of the ratio scale, consider the measurement of response time in a psychology experiment. The ratio scale has an absolute zero (i.e. clock starts ticking at 0 seconds) we can determine the ratio of values.

For example candidate who took 10 seconds to respond to a stimulus can be said to have taken twice as long as a candidate who took five seconds to respond. (A ratio of 2 to 1) The distinction in psychology between interval and ratio measures is unimportant. They are treated as same.

2. Test Measurement

Scale	What do we know	What does it tell us	Examples
Nominal	Equal vs. not equal	The same or not the same	Shoe color
Ordinal	Order	Above or below	Order in a race
Interval	Equality of intervals	The amount above or a certain amount below	Temperature Celsius
Ratio	Equality of ratios	The amount & the proportion above & below	Time (seconds)

3. Test Development

The general process of constructing a test involves determining the purpose of the test, defining the attribute to be measured, developing a measure plan, writing items, conducting a pilot study and item analysis, selecting items, establishing norms and determining the reliability and validity of the measure

3. Test Development

Defining the Purpose

The first step in test development is defining the purpose of the test. This stage provides the foundation for all other development activities.

The test may be designed to access an attribute, predict future performance (e.g. cognitive ability test used to select candidates most likely to succeed in the job), evaluating performance competence (e.g. test for reading ability to assess proficiency), diagnose individual strength and weaknesses (e.g. measure of performance by a supervisor), evaluate programs or provide feedback.

The intended use of the test directs attributes, types of items, length and complexity of the test

3. Test Development

Defining the Attribute

The next step in test development is to define precisely the attribute to be measured.

To clarify the attribute, it is necessary to include a definition of the attribute, the content to be measured (the testing universe), the format for the questions, and instructions for administering and scoring.

A thorough description of the attribute provides a domain of writing items for the test.

3. Test Development

Developing a Measure Plan

After deciding the purpose and defining attributes, nest step is to establish the measure plan. The measure plan must include an outline of content to be included in the test which is derived from the attribute and must cover all the important aspects of the attributes.

It needs to have a brief description of the target group of people who will be taking the test. This helps test developers to write the questions (items) and instructions and the test format (multiple choice, true/false and forced choice) or subjective format (essays & interviews)

3. Test Development

Developing a Measure Plan

Projective tests are another type of subjective format that uses ambiguous stimuli (words or pictures) to prompt responses from the test taker.

Some people have response sets – patterns of responding that result in *false or misleading information* – such as social desirability, acquiescence, random responding and faking that can result in errors in test scores.

Test developer needs to be aware of such responses and guard against them.

3. Test Development

Developing a Measure Plan

Objective items provide ample sampling of the testing universe but they are more time consuming to develop.

Scoring of objective items is easier and likely to be more accurate and reliable. Subjective items are easier to construct and revise.

Some experts suggest that essay and interview questions are better suited for testing higher-order skills such as creativity and organization. Scoring of subjective items is more difficult, requiring independent scoring by two experts to increase reliability and accuracy.

3. Test Development

Writing the Items

Next using the definition of the attribute and the measure plan as the guidelines, items are written. Although the test items make up the bilk of the new test, they are meaningless without specific instructions on how to administer and score the test.

The test developer should write three sets of instructions: one for the administrator, another for the taker and a third for the person who scores and interprets the test results.

3. Test Development

Conducting a Pilot Test

After writing the test items. The test developer conducts a pilot test to determine the final form the test will take. The test developer then follows up the pilot test with other studies that provide the necessary data for validation and norming. Conducting the pilot test and analyzing its data are an integral part of development process.

The pilot test is a scientific investigation of the new test's reliability and validity for its specified purpose.

3. Test Development

Conducting a Pilot Test

Because the purpose of the pilot test is to study how well the test performs, it is important that the test is given in a situation that matches the actual circumstances in which the test will be used.

Test developers may use questionnaire or interviews to gather extra information about the respondents or the test.

Each item in a test is a building block that contributes to the test's outcome or final score.

3. Test Development

Conducting a Pilot Test

Therefore, developers examine the performance of each item

- **to identify those items that perform well,**
- **to revise those that could perform better and**
- **to eliminate those that do not yield desired information.**

3. Test Development

Conducting a Pilot Test

Item characteristic curves provide pictures of each item's difficulty and discrimination.

Developers analyze each item for its

- **difficulty (the percentage of test takers who respond correctly),**
- **discrimination (how well it separates those who show a high degree of the construct from those who show little of the construct)**
- **correlation with other items (for reliability) and with an outside criterion (for evidence of validity)**
- **and bias (whether it is easier for one group than for another group)**

3. Test Development

Revising the Test

Test items are dropped based on their consistency, difficulty, discrimination, and bias until a final form of the test is reached.

If enough items were piloted, no items will be rewritten. Items that are rewritten must be piloted to be sure they now meet criteria for retention in the test.

Other test components such as the instructions, should be revised based on the results of a qualitative analysis in the form of a questionnaire / interviews with all concerned.

3. Test Development

Validation, Norms, Cut Scores and the Test Manual

Standards for validation study are similar to those for designing the pilot study, including using a representative sample of the target audience that is large enough to conduct the desired statistical tests.

Developers conduct the validation study by administering the test to another sample of people.

The validation study provides data on test's reliability, its correlation with any appropriate criteria such as performance evaluations & its correlation with the other measures of the test's construct.

3. Test Development

Validation, Norms, Cut Scores and the Test Manual

If the validation study provides sufficient evidence of reliability and validity, the test developers conduct a final analysis called cross-validation – a final round of test administration to yet another sample of test takers.

After validation is complete, norms and cut scores can be developed from validation data to provide test users with information for interpreting test scores.

Test developers can develop *norms* (distribution of test scores used for interpreting an individual's test score) and *cut scores* (decision points for pass/fail groupings)

3. Test Development

Validation, Norms, Cut Scores and the Test Manual

There are a variety of types of norms, including percentile ranks, age norms and grade norms. Test users should be careful to select the appropriate norm group, ensure that the norms they use are up-to-date and ensure that the size of the norm group is large enough to representative of the population.

NORMING:
Comparing a test score to other relevant past scores.

NORM GROUP
Group whose test scores are used to compare and individual's test score.

3. Test Development

Validation, Norms, Cut Scores and the Test Manual

At the end of the validation process, the test manual is assembled and finalized. Contents of the manual include the rationale for constructing the test, a history of the development process, the results of the validation studies, a description of the appropriate target audience, instructions for administering and scoring the test and information on interpreting individual scores.

Finally, the developers compile the test manual, which was in the making along with the test. It includes answer key, instructions for the administrator and test user, information on test development , validation and cross validation, norms and cut scores

4. Test reliability

Reliability: the term reliability refers to consistency of test results.

Reliability is demonstrated by the consistency of scores obtained when the same applicants are re-examined with the same or equivalent form of an assessment. No assessment procedure is perfectly consistent.

If an applicant's keyboarding skills are measured on two separate occasions, the two scores (e.g. net words per minute) are likely to differ.

Reliability reflects the extent to which these individual score differences are due to "true" differences in the competency being assessed and the extent to which they are due to chance or random errors.

4. Test reliability

A goal of good assessment is to minimize random sources of error. As a general rule the smaller the amount of error, higher the reliability.

Reliability is expressed as a positive decimal number ranging from 0 to 1.00, where 0 means the scores consist entirely of error; and 1.00 means no errors.

For most assessment applications, reliabilities above .70 are likely to be regarded as acceptable. This reliability is important as scores are used to make important decisions.

4. Test reliability

Calculating and Interpreting the Reliability Coefficient

A reliability quotient is calculated by correlating the scores of test takers on two administrations of the same test.

A reliability quotient is interpreted by examining its sign (positive or negative) and its proximity to 100. To be useful, the quotient should be positive and over .70

The standard error of measurement (SEM) provides a measure of how much an individual's score is likely to differ from individual's true score.

Using the SEM, we can calculate a confidence interval that is likely to contain true score.

Cohen's kappa is a measure of agreement for use with nominal or ordinal data such as ratings and decisions. Nearer kappa is to 100, the stronger the agreement.

4. Test reliability

Factors that Influence Reliability

- **Errors that increase or decrease individual scores and change the reliability result from four sources, the test itself, test administration, test scoring and test takers.**
- **Errors that result from poor test design, including ambiguity and poorly worded questions as well as reading levels that are too high for test takers.**
- **Test administration can generate error when administrators do not follow instructions for test administration or when testing environment is uncomfortable or distracting.**

4. Test reliability

Factors that Influence Reliability

- **Tests must be scored accurately and according to the instructions available in the test manual.**
- **Test takers can contribute to test error by being fatigued or ill, by cheating, or by providing dishonest answers.**
- **Reliability is related to test length, the longer the test, the more reliable it is likely to be – provided that the questions on the test are homogenous.**

4. Test reliability

Methods to Determine Reliability of a Test

- **Test – Retest**

Test- retest involves giving the same test to the same people at different times. Scores are correlated from time 01 and time 02 to get a reliability coefficient referred to as coefficient that assesses the amount of error due to random fluctuations in score over time.

Closer the scores are for the individual at different times, the more reliable the test is. But individual scores can improve as a result of practice effect.

4. Test reliability

Methods to Determine Reliability of a Test

★ **Test – Retest :contd.**

To over come practice effects and differences in individuals and the test administration one time to the next, psychologists often give two forms of the same test – alike in every way – to the same people at the same time.

This method is called *alternate or parallel forms.*

4. Test reliability

Methods to Determine Reliability of a Test

- **Alternate or Parallel Forms**

This method examines the consistency with which an attribute is measured across different versions of the test.

This is achieved by calculating the correlation between two forms to obtain a coefficient of equivalency. Two forms can administered close together, but to prevent order effects, half of test takers be given form A first and other half form B.

4. Test reliability

Methods to Determine Reliability of a Test

Internal consistency reliability indicates the extent to which items on test measure the same thing.

★ **Alternate or Parallel Forms: contd.**

Error using this method is defined as *content sampling*. A high parallel form reliability coefficient indicates that the different forms of the test are very similar which means that it makes virtually no difference which version of the test a person takes. On the other hand, a low parallel form reliability coefficient suggests that the different forms are probably NOT comparable ; they may be measuring different things and cannot be used interchangeably.

4. Test reliability

Methods to Determine Reliability of a Test

★ Internal Consistency

Internal Consistency determines the degree to which the various items of the test correlate with each other.

Error is defined as item heterogeneity: the more homogenous the items lower the error. When items are correlated, it shows they are measuring the same attribute.

"Internal consistency reliability indicates the extent to which items on a test measure the same thing"

4. Test reliability

Methods to Determine Reliability of a Test

★ Inter – Rater

Inter- rater reliability is a method of estimating reliability or accuracy of those making judgments when scoring a test. Two types of scorer reliability are important: the comparison of judgments between or among scorers and the comparison of judgments each scorer makes for all tests. We use correlation to calculate the inter-rater reliability for interval data such as numerical test scores.

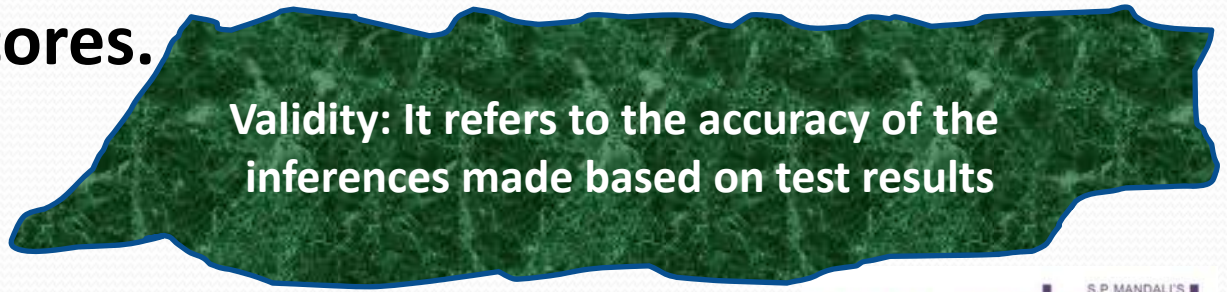
"Inter- rater reliability is a method of estimating reliability or accuracy of those making judgments when scoring a test.

5. Test 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).

Validity refers to the relationship between performance on an assessment and performance on the job.

It provides evidence attesting to what attribute a test is measuring, how well it is measured and what decision can be made from the scores.



5. Test Validity

Thus the process of validation evaluates if a test is assessing the attribute it is supposed to and if a measure can be used to make accurate decisions.

The test itself is not validated but the inferences of the test are validated. Investigations are conducted to gather evidence to support theses inferences.

Evidence is continually gathered to evaluate and revise a test if it is not fulfilling the intended purpose. Therefore, validation is an on-going process.

5. Test Validity

This evidence might be

- **documentation of links between the test and requirements of the job or**
- **showing that the test relates to other measures of the same thing.**

Test publishers have information on the validity of their test products.

Factors that outline the judgment regarding what types of validity evidence is valid for a given test is provided in the Standards for Educational & Psychological Testing .

5. Test Validity

At this stage a distinction needs to be made between validity and reliability.

Suppose a clock is set forward by 30 minutes. If the time piece is good, time it shows will be reliable (consistently 30 minutes forward), but will not be valid as judged by 'standard time,'

The reliability is found by making reported measurements of the same facts while validity is found by comparing the data obtained from the test with the standard measures.

- ***A test is valid for a particular purpose, it is not generally valid.***

5. Test Validity

There are many forms of validity evidence.

Which type is most appropriate depends on how the assessment is used in making an employment decision. For example, if a work sample test is designed to mimic the actual tasks performed on the job, then a content validity approach may be needed to establish the content of the test matches in a convincing way the content of the job, as identified by the job analysis.

When multiple selection tools are used, one can consider combined validity of the tools. Used together, the tools can more accurately predict the applicant's job performance than any tool used alone.

5. Test Validity

The amount of predictive validity one tool adds relative another is often referred to as the incremental validity of the tool.

The incremental validity of an assessment is important to know because even if an assessment has low validity by itself, it has the potential to add significantly to the prediction of job performance when joined with another measure.

Just assessment tools differ with respect to reliability, they also differ with respect to validity.

The next table shows estimated validities of various assessment tools for predicting job performance as well as the incremental validity gained from combining each with a test of cognitive ability.

5. Test Validity

Validity of various Assessment Tools alone and in combination			
Assessment Method/Test	Validity if used alone	Incremental validity	**% increase due to combining with *(CA)***
Cognitive Ability (CA)	.51		
Work Sample	.54	.63	**24**
Structured Interviews	.51	.63	**24**
Job Knowledge	.48	.58	**14**
Accomplishment Record	.45	.58	**14**
Integrity/honesty	.41	.65	**27**
Unstructured Interviews	.38	.55	**8**
Assessment centers	.37	.53	**4**

5. Test Validity

Validity of various Assessment Tools alone and in combination contd.

Assessment Method/Test	Validity if used alone	Incremental validity	% increase due to combining with (CA)
Bio data Measures	.35	.52	**2**
Conscientiousness	.31	.60	**18**
Reference checking	.26	.57	**12**
Job experience	.18	.54	**6**
Training & Experience Point	.11	.52	**2**
Years of Education	.10	.52	**2**
Interest	.10	.52	**2**

Using Validity Information and Ethical Issues

When a relationship can be established between a test and a criterion, the test scores can be used to predict how well individuals are likely to perform on the criterion.

Decisions based on test scores have far-reaching consequences and test users should have evidence of validity to be sure that the test scores and their predictions will be appropriate.

Using Validity Information and Ethical Issues

Researchers, test developers, publishers and test users are ethically and morally responsible for ascertaining that any psychological test used for making predictions and decisions shows acceptable evidence of reliability and validity.

Their responsibility also extends to guarding against test misuse and to increasing public awareness about the important role that tests play in test takers lives.

Three major types of validity evidence are discussed next. They are inter-related and are used in combination also.

5.1 Criterion Related Validity

CRV is particularly suited to determine if a test can be used to make predictions and /or decisions. It involves co-relating scores on a predictor (measure of test) with some criterion (level of success) to determine if accurate decisions can be made from the scores.

For example, a fruit grower might want to determine how valid – as predictor of grafting ability – is a manual dexterity test in which farm workers have to quickly arrange wooden pegs in a box. If a substantial statistical relationship exists between performance on the test and in the field, the grower might want to use the test to hire grafters – who will never deal with wooden pegs in the real job.

5.1 Criterion Related Validity

CRV : This measurement compares a test's results to the individual's behavior in the real world.

Correlations may range from -1 to 0 to a +1.

In linear (positive) relationship applicants who did well on a test would do well on the job, those who did poorly on the test would do poorly on the job.

In a negative (or inverse) relationship applicants who did well on a test would do poorly on the job, those who did poorly on the test would do well on the job.

A correlation coefficient quotient close to "0", would indicate that test and performance are not related in any way.

5.1 Criterion Related Validity

Expect correlation coefficients that measure reliability to be higher than those convey validity.

Subjective meanings for validity coefficients.

Correlation Coefficient	Subjective meaning for validity
r = .40 or greater	Somewhat acceptable.
r = .50 or greater	Good
r = .60 or greater	Excellent

5.1 Criterion Related Validity

There are two types of criterion related validity:

5.1.1 Concurrent Validity	
Refers to how well test matches up with an individual's current performance in an area of interest	
5.1.2 Predictive Validity	
Refers to a test's ability to predict an individual's future performance in a given area.	

5.1.1 Concurrent Validity

Concurrent Validity evaluates if an individual's level of an attribute is acceptable to achieve the criterion at the present time. In a concurrent study, job incumbents (present employees) are tested and their job performance is evaluated at the same time.

The correlation between current performance on assessment and on the job then can be examined. Evidence of concurrent validity is often substituted for predictive validity. Whether this is appropriate will depend upon type of measure and how similar the incumbent sample is to the applicant population.

5.1.2 Predictive Validity

In a predictive validation, job applicants are tested and their job performance is evaluated at a later time, usually after being on the job for six months or more. The relation between performance on the assessment and on the job then can be examined.

Predictive studies begin with obtaining scores from random sample of the population in which decisions are made without using scores from the test.

After the decision is made, scores from the criterion are gathered and correlation between the test and the criterion is calculated.

5.1.2 Predictive Validity

An example of this is when job applicants have been hired and worked for a period of time , information on absenteeism, thefts, and other counter-productive behaviors are collected and correlated with the integrity test to define its predictive ability.

5.2 Content Related Validity
Evidence is based on job analysis and expert judgment and the choice of items or tasks included in the assessment logically match or represent those tasks or competencies required by the job.
5.3 Construct Related Validity
Construct validation involves collecting evidence to determine whether the assessment does indeed measure the trait it was intended to measure.
5.4 Face Validity
The face validity of a test is perhaps the simplest measure of a test's validity. It refers to whether or not the test's material has any relevance to what the test is supposed to measure.

5.2 Content Related Validity

In *Content Related Validity*, the content of the job is clearly mirrored in the selection process. This approach is useful to the degree that the selection process and the job are related. Thus it makes sense for clerk-typist to be given a typing test and so on.

The pitfall of this method is that people tend to examine only those areas that are easiest to measure. If important skills of the job are not tested (as they are difficult to measure), the approach becomes ineffective.

Methods

1. Perform a series systematic steps during test development to ensure that the test samples the construct being measured representatively.

2. Review test items after the test development and determine the extent to which experts agree that test items are essential.

5.3 Construct Related Validity

In *Construct Related Validity* is the third type of evidence that can be used to decide if inferences made from a test scores are valid. Construct is the process of accumulating evidence to establish if the test is assessing the attribute it is intended to assess.

However, instead of evaluating the test plan and the content of the attribute, construct related validity investigates hypothesized relationships between a construct and other constructs to assess if actual relationships are similar to predicted ones.

5.3 Construct Related Validity

This type of validity refers to how well a test's results correspond to the individual's characteristics in the area being measured. The construct validity of a test is difficult to determine but it can be done by comparing a test's results with that of other tests or by comparing an individual's scores before and after some sort of training designed to alter the characteristics being tested.

The process of gathering evidence covers three steps like

5.3 Construct Related Validity

i. defining the construct

ii. identifying observable behaviors that define the construct

ii. checking if observable behaviors are good indicators of a construct

If so, they are regarded as highly correlated indicating that they are measuring the same concept.

Once this internal consistency of behaviors is established, a logical network is constructed. This network specifies variable to which the construct should or should not be related. Closer the match, stronger the evidence of construct validity.

5.4 Face Validity

Face Validity refers to what a selection process appears to measure on the surface. Although face validity is not a type of validation strategy, it is usually vital that a selection approach appear to be valid, especially to the applicant. A farmer wanting to test for a herdsman's knowledge of math should use test problems involving dairy matters, rather than questions using oranges and apples. The skills could be determined by either approach, but applicants often , rightfully, resent being asked questions they feel are not related to the prospective job.

5.4 Face Validity

Face Validity is a desirable attribute of a selection process. Not only does it contribute toward a realistic job preview, it also helps eliminate negative feelings about the process.

Though face validity alone cannot be used to support the use of an assessment, it is important because it promotes cooperation and acceptance of the process on the part of applicants. Face validity is the simplest measure of a test's validity as its only objective is to verify that test material has relevance to what it is designed to measure.

6. The Types of Psychological Testing

Tests can be classified as

- **Maximal performance**
- **Behavior observation or self report**
- **Standardized or non-standardized**
- **Objective or projective**
- **By dimension measured or by subject.**

6. The Types of Psychological Testing

Differences in tests are caused by

- **Behavior test takers perform**
- **Attribute, trait or characteristic measured**
- **Outcome predicted**
- **Content, format, and how administered**
- **How scored and interpreted**
- **Psychometric quality**

6.1 Intelligence Test

6.1.1 The Stanford-Binet Intelligence Scale

At the beginning of 19th century Alfred Binet began designing an intelligence test that emphasized verbal and was arranged in order of difficulty. Items were grouped by level so that 80 – 90% of any given age could pass that age's level.

The basis for Binet's evaluation was the assumption that people who performed below their age level were retarded, those who performed at it were normal and those who performed above it were gifted.

6.1 Intelligence Test

6.1.1 The Stanford-Binet Intelligence Scale

Lewis Terman revised Binet's test for use in the US, when he was at Stanford University . The new test is known as Stanford-Binet test.

The new test was individually administered and the first to use the idea of an intelligence quotient or IQ, which could be derived from a ratio between a person's mental age (MA) and chronological age (CA)

$$IQ = (MA/CA) \times 100$$

6.1 Intelligence Test

6.1.1 The Stanford-Binet Intelligence Scale

The Stanford-Binet test, edition 5, SB5 has now five factors Fluid Reasoning, Knowledge, Quantitative Reasoning, Visual-Spatial Processing and Working Memory.

Unique to SB5 is the use of non-verbal mode of testing covering all five cognitive factors. The range of the scales has been extended to more accurately measure both higher and lower areas of functioning.

6.1.1 The Stanford-Binet Intelligence Scale

Scoring

Raw scores are converted into scaled scores (M = 10, SD = 3) using age appropriate tables. The scaled scores are summed for Nonverbal, Verbal and Full Scale IQ as well as for the five factor index scores. These are all normalized standard scores (M = 10, SD = 3). Percentile rank equivalent and the confidence intervals are also obtained.

Reliability

Using the split half method and correcting with the Spearman-Brown formula, reliability coefficients were very high for the Full Scale Score. The non-verbal and Verbal showed excellent stability.

6.1.1 The Stanford-Binet Intelligence Scale

Validity

As with the SB:FE, several studies were done to investigate the validity of the SB5. comparisons were also made with other tests. High correlation between the composite IQ scores of the SB5 and the scores of previous SB edition and all of the major IQ batteries used for all populations.

SB test is good because it has become very well standardized. The test is also one of the very few which can be administered to someone several times at a variety of ages.

6.1.2 The Wescher Adult Intelligence Scales

The WAIS is a test designed to measure intelligence in adults and older adolescents. It is currently in its fourth edition. Tests were innovative in 1930s because they gathered tasks created for nonclinical purposes for administration as a "clinical test battery".

Because Wescher tests included non-verbal items (known as performance scales) as well as verbal items for all test seekers. WAIS-4 was released in 2008 and is composed of 10 core subtests and five supplemental subtests, with 10 core subtests comprising the full scale IQ.

6.1.2 The Wescher Adult Intelligence Scales

Wescher defined intelligence as " … the global capacity of a person to act purposefully, to think rationally and to deal effectively with his environment."

Indices and Scales

There are four index scores representing major components of intelligence:

- **Verbal Comprehension Index (VCI)**
- **Perceptual Reasoning Index (PRI)**
- **Working Memory Index (WMI)**
- **Processing Speed Index (PSI)**

6.1.2 The Wescher Adult Intelligence Scales

Indices and Scales

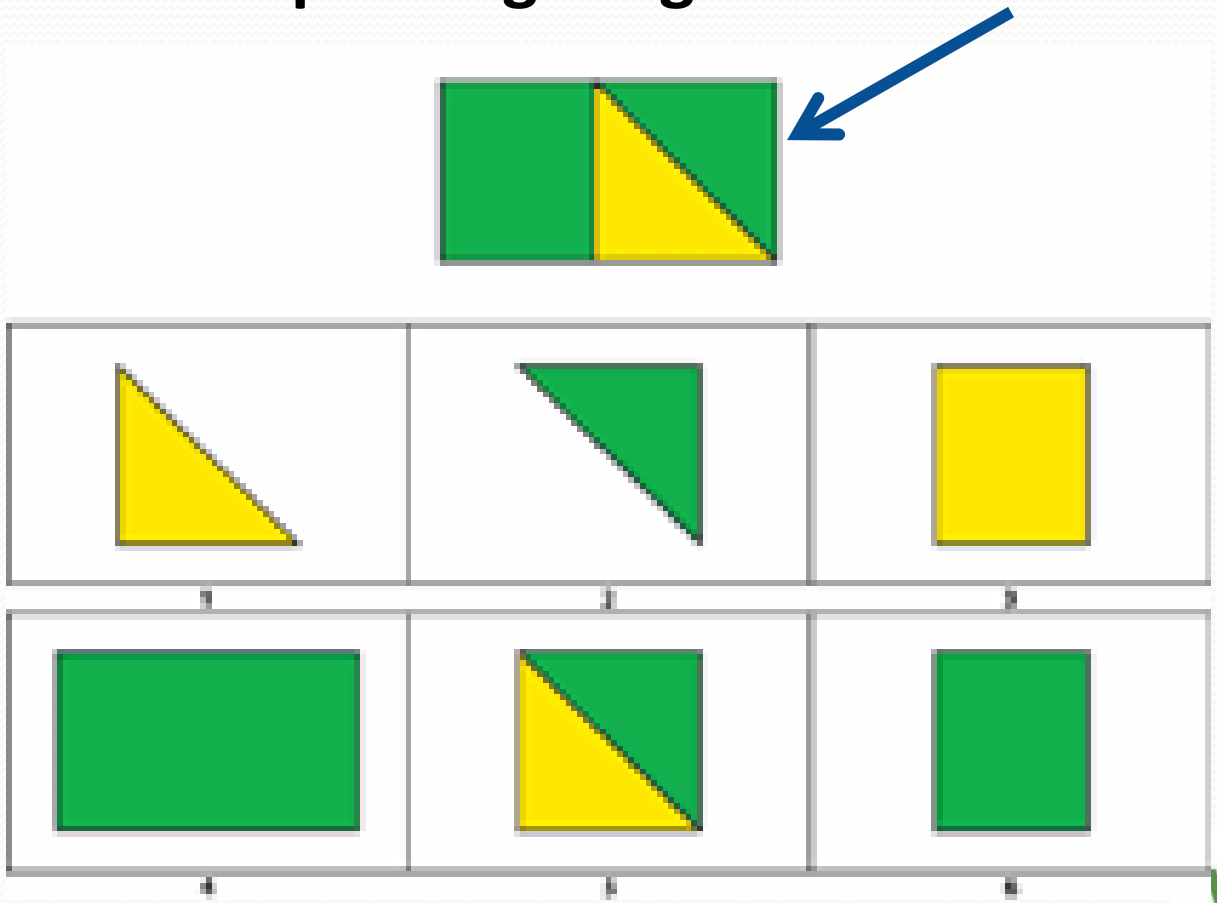
Two broad scores are also generated, which are used to summarize general intellectual abilities:

- **Full Scale IQ (FSIQ) based on total combined performance of VCI, PRI, WMI and PSI.**
- **General Ability Index (GAI) based only on six subtests that the VCI and PRI comprise.**

6.1.2 The Wescher Adult Intelligence Scales

Activity 01

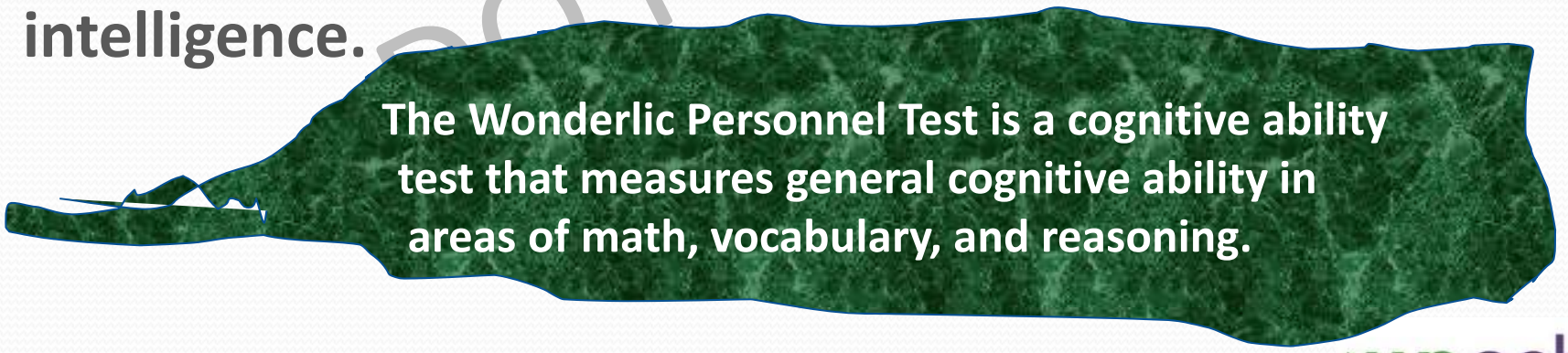
Which 3 of these pieces go together to make this puzzle?



6.1.3 Wonderlic Cognitive Ability Test

The Wonderlic Cognitive Ability Test is a popular group intelligence test used to assess the aptitude of prospective employees for learning and problem solving in a range of occupations. It consists of 50 multiple choice questions to be answered in 12 minutes.

The test was developed by Eldon F Wonderlic. The score is calculated as the number of correct answers given in the allotted time. A score of 20 is intended to indicate average intelligence.



6.1.3 Wonderlic Cognitive Ability Test

Created in 1936, the Wonderlic personnel test is a cognitive ability test that measures general cognitive ability in the areas of math, vocabulary and reasoning. There are currently 30 tests offered and they present questions in an open response format with increasing difficulty.

The tests are divided into four different sections cognitive, skill, personality and behavioral. The scores are predictors of the possible conformity a potential employee has within the field that they are applying. These measures of general mental ability are used for employee selection, training and placement.

6.1.3 Wonderlic Cognitive Ability Test

Activity 02

See how you score on some examples from a Wonderlic IQ test. Set your clock for five minutes, don't peak at the answers. And …oh, yeah, run the 40 and give us some bench-presses first would you?

1. **Look at the numbers below. What number should come next?**

8	4	2	1	½	¼	?

2. **Assume the first two statements are true. Is the final one: True False Not Certain ?**

The boy plays baseball. All baseball players wear hats. The boy wears a hat.

6.1.3 Wonderlic Cognitive Ability Test

Activity 02

3. Paper sells for 21 cents per pad. What will four pads cost?

4. How many of the five pairs of items listed below are exact duplicates?

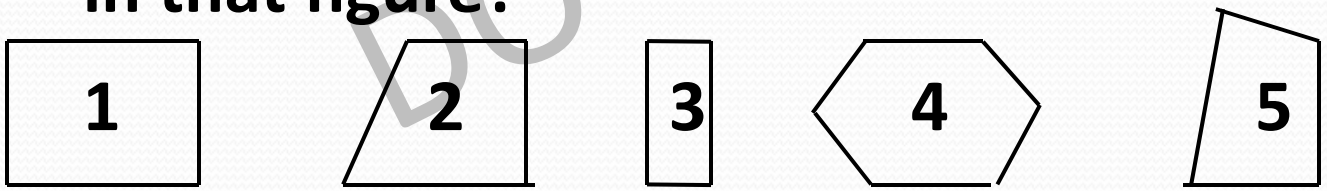
Nieman, K M	Neiman, K M
Thomas, G K	Thomas, C K
Hoff, J P	Hoff, J P
Pino, L R	Pina, L R
Warner, T S	Wanner, T S

6.1.3 Wonderlic Cognitive Ability Test

Activity 02

5. RESENT RESERVE
   1. have similar meanings
   2. have contradictory meanings
   3. mean neither the same nor the opposite

6. One of the numbered figures in the following drawing is most different from the others. What is the number in that figure?



6.1.3 Wonderlic Cognitive Ability Test

Activity 02

7. A train travels 20 feet in 1/5 second. At the same speed, how many feet will it travel in three seconds?

8. When rope is selling at $ .10 a foot, how many feet can you buy foe sixty cents?

9. The ninth month of the year is
   1. October 2. January 3. June 4. September 5. May

10. Which number in the following group of numbers represents smallest amount?
    7 8 31 .33 2

6.1.3 Wonderlic Cognitive Ability Test

Activity 02

11. In printing an article of 48,000 words, a printer decides to use two sizes of type. Using the larger type, a printed page contains 1,800 words. Using smaller types, a page contains 2,400 words. The article is assigned 21 full pages in the magazine. How many pages must be in smaller type?

12. The hours of daylight and darkness in SEPTEMBER are nearest equal to hours of daylight and darkness in:

1. June 2.March 3.May 4. November

6.1.3 Wonderlic Cognitive Ability Test

Activity 02

13. Three individuals form a partnership and agree to divide the profits equally. X invests S 9,000, Y invests $ 7,000 , Z invests $ 4,800, how much less does X receive than if the profits were divided in proportion to the amount invested?

14. Assume the first two statements are true. Is the final one: 1.True 2. False 3. Not certain ?

    Tom greeted Beth. Beth Greeted Dawn. Tom did not greet Dawn.

15. A boy is 17 years old and his sister is twice as old. When the boy is 23 years old, what will be the age of the sister?

6.1.3 Wonderlic Cognitive Ability Test

Activity 02

Answers:

1. ⅛	6. 4	11. 17
2. True	7. 300 feet	12. March
3. 84 cents	8. 6 feet	13. $ 560
4. 1	9. September	14. Not certain
5. 3	10. .33	15. 40 years old

6.1.4 The Chicago Test

Louis Leon Thurstone was a pioneer in the fields of psychometrics and psychophysics. He developed a new factor of analytic techniques to determine the number and nature of latent constructs within a set of observed variables.

The Chicago Test of primary Mental Abilities, a battery of 56 tests that were administered to 240 volunteers, most of whom were students at the University of Chicago. Instead of viewing intelligence as a single, general ability, Thurstone's theory focused on seven different "primary mental abilities."

6.1.4 The Chicago Test

The abilities Thurstone described were:

Verbal comprehension – the ability to define and understand words.
Word fluency – the ability to produce words rapidly.
Number – the ability to solve arithmetic problems.
Space – the ability to visualize relationships.
Memory – the ability to memorize and recall.
Perception - the ability to see differences and similarities among objects.
Reasoning – the ability to find rules

6.1.4 The Chicago Test

The Chicago Test is focused on seven different primary mental abilities instead of viewing intelligence as a single , general ability

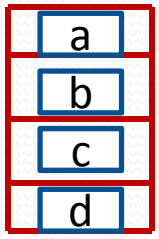
After deciding that these seven factors made up the intelligence, Thurstone rearranged the existing tests and devised some new ones.

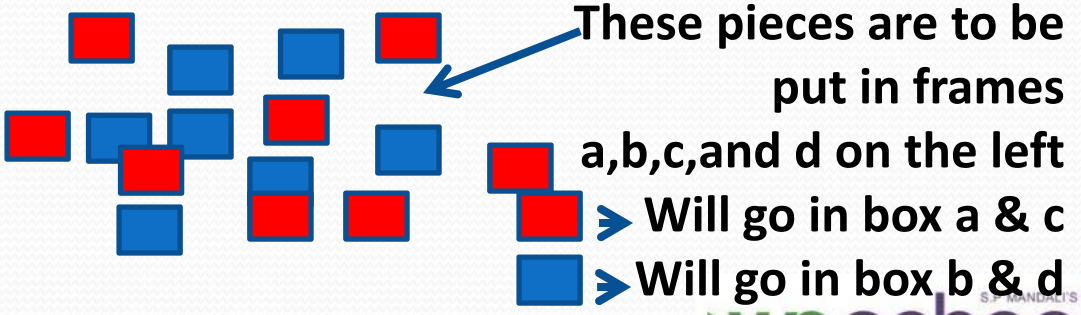
He established that intelligence consists of an independent primary factor and a general factor that is shared by all primary factors. Intelligence consists of both general ability and a number of specific abilities.

6.2.1 The Minnesota Spatial Relations Test

This test contains geometric pieces which fit into their matching frames on two different boards. The examinee's job is to pick the frame which correctly represents the assembled geometric pieces. Test scores correlate positively with grades in shop and engineering courses.

Test scores also correlate positively with supervisor ratings as well as performance records in inspection, packing, machine operation in industrial settings.





6.2 Ability Test

6.2.1 The Minnesota Spatial Relations Test

Another descendant of the MSRT is the revised Minnesota Spatial paper Form board test (Likert & Quasha, 1995)

This paper and pencil adaptation of MSRT was designed for grades 9 – 16 and adults. It consists of 64 multiple choice items, each containing a frame showing geometric divided into several parts and five answer frames containing an assembled form. The candidate's task is to select the one answer frame out of five showing how the disassembled geometric figure will look if the parts are joined together.

6.2 Ability Test

6.2.2 The Bennett Mechanical Comprehension Test

The Bennett Mechanical Comprehension Test is assessment that measures an individual's aptitude to learn mechanical skills. It measures a complex set of abilities composed of three primary facets or constructs.

i. **Mechanical information.**
ii. **Spatial visualization**
iii. **Mechanical reasoning or understanding**

The knowledge measured by this test is based on common experiences and special training has little effect on test scores.

6.2 Ability Test

6.2.2 The Bennett Mechanical Comprehension Test

The BMCT has 68 multiple choice items that present simple, frequently encountered mechanisms and situations. While not based on specific training, BMCT does require a working knowledge of basic mechanical operations and the application of physical laws.

It is a timed test to be completed in 30 minutes or less. The BMCT focuses on spatial perception and tool knowledge rather than manual dexterity. It is especially well suited for assessing job candidates for positions that require a grasp of the principles underlying the operation and repair of complex devices.

6.2 Ability Test

6.2.2 The Bennett Mechanical Comprehension Test

- is an assessment that measures an individual's aptitude to learn mechanical skills.

The BMCT helps to quickly and effectively :

- ✓ Identify candidates with good spatial perception and mechanical reasoning ability.
- ✓ Assess a candidate's knowledge of basic mechanical operations and physical laws.
- ✓ Recognize an aptitude for learning mechanical processes and tasks.
- ✓ Predict employee success and appropriately align your workforce.

6.2 Ability Test

6.2.2 The Bennett Mechanical Comprehension Test

Who is BMCT for?

- ✓ **Automotive or Aircraft Mechanic**
- ✓ **Engineer**
- ✓ **Installation / Maintenance / Repair**
- ✓ **Industrial / Technical Sales**
- ✓ **Skilled Tradesperson (e.g. Electrician, Welder, Carpenter)**
- ✓ **Transportation Trades / Equipment Operator**

6.2 Ability Test

6.2.2 The Bennett Mechanical Comprehension Test

Scoring

The BMCT is scored by hand with a scoring template. The raw score is simply the number correct. Raw scores are converted to percentile scores based on the appropriate normative population. Scoring instructions for BMCT are clear and easy to master.

Reliability

The manual of the BMCT presents split-half reliability coefficient of .81 to .93, with a median of .86. the standard errors of measurement range from 3.0 to 3.8.

6.2 Ability Test

6.2.2 The Bennett Mechanical Comprehension Test

Validity

Current validity is not discussed in the manual because there is no specific job or curriculum against which to compare test items. Construct validity is also not specifically addressed. A table of correlations of BMCT with other standardized tests from an industrial data base is represented.

Norms

To develop Forms S and T, 180 experimental items were administered to 706 male students.

6.2 Ability Test

6.2.3 The Detroit Clerical Test

- is designed to identify candidates who can work quickly and accurately in administrative roles.

Because mistakes in accounting, billing or shipping information can be very expensive, organizations are increasingly using this type of test when recruiting such positions as clerical jobs, bank staff, cashiers, warehouse workers and other positions requiring concentrated work with client details.

There are two types of test that are specific to clerical and administrative roles.

1. Concentration/ Work Rate
2. Data Checking

6.2 Ability Test

6.2.3 The Detroit Clerical Test

1 Concentration/ Work Rate Tests

Concentration tests are speed tests. This means that given sufficient time to complete them, most people would be able to obtain perfect score, provided that they were capable of working in a systematic and careful way. However, the time limit is usually set so that test is impossible to complete, the questions tend to be similar and rather repetitive which makes it difficult to maintain attentiveness. Therefore, concentration tests are *to select personnel who needs to work through items of information in a systematic way while making very few mistakes.*

6.2 Ability Test

6.2.3 The Detroit Clerical Test

1 Concentration/ Work Rate Tests

The question comprises of :

1. **A grid in which there are rows of symbols, numbers and letters.**
2. **A key which consists one symbol, one number and one letter.**
3. **A series of five answer options.**

In the grid below the letter L has been taken from column 2, the symbol ₩ has been taken from column 3 and number 4 has been taken from column 5.

6.2 Ability Test

6.2.3 The Detroit Clerical Test

1 Concentration/ Work Rate Tests

∞	↔	₩	Δ	₪	Ɗ
3	7	6	1	4	8
E	L	K	C	G	P

KEY	
	L
	₩
	4

Δ 1 G	₪ P 7	6 ↔ G	C 7 ↔	7 Ɗ 6
A	B	C	D	E

6.2 Ability Test

6.2.3 The Detroit Clerical Test

1. **Concentration/ Work Rate Tests**

The candidate has to examine answer options and decide – in which option have the three elements have been taken from the same columns as in the key. In other words you are looking for the option where one element has been taken from column 2, one from column 3 and one from column 5.

6.2 Ability Test

6.2.3 The Detroit Clerical Test

2 **Data Checking Tests** - are used to select candidates for clerical and data input jobs, particularly where accuracy is important.

Data Checking Tests present the candidate with number of tables of information which must be checked against each other. This type of test is used measure how quickly and accurately errors can be detected in the data. These tests usually contain between 20 to 40 questions and take 10 – 20 minutes to complete. It is important, although difficult, to maintain concentration for the full duration of the test.

The speed at which questions are answered is the critical measure.

6.2.4 The Reaction Time Test

Reaction time is the time elapsed between the presentation of sensory stimulus and the subsequent behavioral response. In psychometric psychology it is considered to be and index of speed of processing. That is, it indicates how fast the thinker can execute the mental operations needed by the task at hand.

In turn speed of processing is considered an index of processing efficiency. The behavioral response is typically button press but can also be an eye movement, a vocal response or some other observable behavior.

The Reaction Time Test indicates how fast the thinker can execute the mental operations needed by the task at hand.

6.2.5 The Manual Dexterity Test

The Manual Dexterity Tests measure a candidate's dexterity and their sense of space and shape relationships. It is a valuable indicator for manual tasks in relation to machine operation. It evaluates manual speed and hand/eye coordination

Created by Joseph Tiffin, the Purdue Pegboard is a manual test consisting of a pegboard, pins, collars and washers. The Purdue Pegboard Test has been used extensively in the selection of employees for jobs that require fine and gross motor dexterity / coordination. It measures gross movements of hands, fingers, arms and fingertip dexterity.

It is intended for industrial use and assembly work in a factory setting.

6.2.5 The Manual Dexterity Test

The Hand Tool Dexterity Test

This test measures proficiency in using ordinary mechanical tools. The test consists of tools and two uprights with bolts, washers and nuts.

The object is to disassemble all the bolts from one upright and reassemble them on corresponding rows of the other upright with heads of the bolts inside.

This type of skill is important to many industrial jobs and apprentice training. Results are used to determine vocational interest and as an indicator of success in similar jobs.

6.3 Aptitude Test

6.3.1 The Differential Aptitude Test (DAT)

The Differential Aptitude Test (DAT) is a multiple aptitude battery designed to measure adult's ability to learn or succeed in certain areas. The test follows the theory that different individuals have varying levels of interest and intelligence in different fields.

One may be good at maths but bad in verbal reasoning. Some may use language excellently but may be very bad in calculations. This variety of aptitudes generates a balance.

6.3 Aptitude Test

6.3.1 The Differential Aptitude Test (DAT)

The Differential Aptitude Test (DAT) contains two levels, with two equivalent alternate forms for each level. Eight subtests measure abilities. A score is provided for each subtest as well as for scholastic aptitude. Test administration procedures are easy to follow and test materials are durable and reusable.

The Differential Aptitude Test (DAT) is a multiple aptitude battery designed to measure adult's ability to learn or succeed in certain areas.

6.3 Aptitude Test

6.3.1 The Differential Aptitude Test (DAT)

A number of tests are available that measure

Verbal reasoning
Numerical ability
Abstract reasoning
Mechanical reasoning
Space relations and
Language usage.

Read more :

http://www.shrm.org/templatesTools/AssessmentResources/SHRMTestingcenter/products/PsychCorp/pages/DAT.aspx

6.3 Aptitude Test

6.3.2 The General Aptitude Test Battery – GATB

GATB developed by the United States Employment Services has proved occupationally very significant. It consists of 12 tests which measure nine aptitudes important for success in a wide variety of occupations.

Eight of the tests are paper pencil tests and are issued in three booklets. The other four are apparatus tests.

Different jobs may require different capabilities to perform them perfectly. Depending upon the job requirements, these tests may either be used individually or in a composite way.

6.3 Aptitude Test

6.3.2 The General Aptitude Test Battery – GATB

GATB is "a series of timed tests used mainly by employment agencies to assist in determining an individual's occupational aptitudes."

The factors covered by GATB are as under:

G Intelligence General learning ability, ability to grasp instructions and underlined principles. It is often referred to as scholastic aptitude and is measured adding scores of three tests also used to measure other factors (vocabulary, arithmetic reasoning and three dimensional spaces)

6.3 Aptitude Test

6.3.2 The General Aptitude Test Battery – GATB

V. Verbal aptitude: ability to understand the meaning of words and paragraph, to grasp a concept presented in verbal form and to present ideas clearly. It is measured by a vocabulary test requiring the subject to indicate which two words in each set have the same or opposite meaning.

N. Numerical Aptitude: ability to perform arithmetic operation quickly and accurately. It is measured by Computation and Arithmetic Reasoning Test. The computation test involves speed and accuracy in simple computation with whole numbers.

6.3 Aptitude Test

6.3.2 The General Aptitude Test Battery – GATB

S. Spatial Aptitude: ability to visualize object in space and to understand the relationship between plans and solid forms. It is measured by Three Dimension Space Test, involving the ability comprehend two dimensional objects and to visualize effects of movements in three dimensions.

P. Form Perception: ability to perceive pertinent detail in objects or in graphic material, to make visual comparisons and discriminations in shapes and shadings. It is measured by two tests requiring the subject to match identical drawings of tools in one test and geographic forms in the other.

6.3 Aptitude Test

6.3.2 The General Aptitude Test Battery – GATB

Q. Clerical Perception: ability to perceive pertinent detail in verbal or numerical material to observe difference in coping tables, lists etc. It might also be called proof reading. It is similar to P but requires matching of names.

K. Motor Coordination: ability to coordinate hand movement with judgments made visually with speed and precision. It is measured by simple paper pencil test requiring subjects to make specified pencil marks in a series of squares.

6.3 Aptitude Test

6.3.2 The General Aptitude Test Battery – GATB

F. Finger Dexterity: ability to move the finger and to manipulate small objects rapidly and accurately. It is measured by two tests requiring assembling and disassembling respectively of rivets and washers.

M. Manual Dexterity: ability to move the hands easily and skillfully. Involves a gross type of movement than finger dexterity. It is measured by two tests requiring subjects to transfer and reverse pages on a board.

The validity of GATB is generalized, meaning that same test score may be predictive of aptitudes for all jobs within a certain skill set.

6.3 Aptitude Test

6.3.3 The SRA Mechanical Aptitudes Test

Mechanical aptitude tests mainly measure potential employee's knowledge of straight forward mechanical and physical concepts. Test scores significantly depend on candidates' mechanical reasoning abilities, including their knowledge of shop arithmetic, levers, pulleys, simple electrical circuits and tools.

Many mechanical comprehension tests are industry specific. The tests apply to positions as welding, plumbing, hydraulics, map reading and auto repair.

6.3.3 The SRA Mechanical Aptitudes Test

The SRA Mechanical Aptitudes Test places more weight on mechanics than on academic depth and generally applies to job applicants for such positions as welding, plumbing, hydraulics, map reading and auto repair.

6.4 Personality Tests

6.4.1 The California Psychological Inventory – CPI

CPI is a self report created by Harrison Gough and currently published by Consulting Psychologists Press. It was created to assess the everyday "folk concepts" that ordinary people use to describe the behavior of people around them.

The CPI is made up of 434 true-false questions. The test is scored on 18 scales , three of which validity scales. Eleven of the non-validity scales were selected by comparing responses from various groups of people. The other four were content validated.

6.4 Personality Tests

6.4.1 The California Psychological Inventory – CPI

It takes about 45 – 60 minutes to complete.

The revised third edition of the CPI contains 434 items and requires that candidates' false and true answers be transformed into raw scale and standard scores by the publisher, who will provide interpretive report writing.

Although it is untimed, the test typically takes an hour to complete and can be done in a group or individual setting.

6.4 Personality Tests

6.4.1 The California Psychological Inventory – CPI

The California Psychological Inventory – CPI is an in-depth personality assessment , measuring personal and professional characteristics, motivations and thinking styles, self management and dealing with others.

6.4 Personality Tests

6.4.1 The California Psychological Inventory – CPI

Scoring:

The inventory contains 434 items which can be scored to yield 18 scales. The 18 scales are further grouped into four classes:

- **Measures of poise, ascendancy, self-assurance and interpersonal adequacy**
- **Measures of socialization, responsibility, intrapersonal values and character.**
- **Measures of achievement potential and intellectual efficiency.**
- **Measures of intellectual modes and interest modes.**

6.4 Personality Tests

6.4.1 The California Psychological Inventory – CPI

Validity :

Correlations between CPI scales and related external criteria tend to fall in the .2 to .5 ranges.

This degree of correlation is typical for much of personality research.

Extremely high correlations are not likely to be found for personality measures because the scales typically try to assess rather broad behaviors.

6.4.2 The Minnesota Multiphasic Personality Inventory

The Minnesota Multiphasic Personality Inventory (MMPI) was developed in the late 1930's by psychologist Starke R Hathaway and psychiatrist J C McKinley at the university of Minnesota. The MMPI has been considered the gold standard in personality testing ever since its inception in 1939. Many additions and changes to the measure have been made over the time.

The MMPI-2 can be used to assess psychological stability in workers in 'high-risk' professions such as airline pilots, police or workers in the nuclear power industry.

6.4.2 The Minnesota Multiphasic Personality Inventory

10 Scales of MMPI

No	Abbreviation	Description	What is measured	Items
1	Hs	Hypochondriasis	Concern with bodily symptoms	32
2	D	Depression	Depressive symptoms	57
3	Hy	Hysteria	Awareness of problems and vulnerabilities	60
4	Pd	Psychopathic Deviate	Conflict, struggle, anger, respect for society's rules.	50
5	MF	Masculinity / Femininity	Stereotypical masculine or feminine interest / behaviors	

6.4.2 The Minnesota Multiphasic Personality Inventory

10 Scales of MMPI – contd.

No	Abbreviation	Description	What is measured	Items
6	Pa	Paranoia	Level of trust, suspiciousness, sensitivity	40
7	Pt	Psychasthenia	Worry, anxiety, tension, doubts, obsessiveness	48
8	Sc	Schizophrenia	Odd thinking and social alienation	78
9	Ma	Hypomania	Level of excitability	46
10	Si	Social Introversion	People orientation	69

6.4.2 The Minnesota Multiphasic Personality Inventory

Example questions:

"I like to visit places I've never before" false (F)

"It would be better if almost all laws were thrown away" (T)

Validity Scales of MMPI-2

The L scale: **Also referred to as the "lie scale". This validity scale was developed to detect attempts by candidates to present them in a favorable light.**

6.4.2 The Minnesota Multiphasic Personality Inventory

The F scale: **this scale is used to detect attempts at "faking good" or "faking bad". Questions are asked to determine if test takers are contradicting themselves in their responses.**

The K Scale: **sometimes referred to as the "defensiveness scale". This scale is a more effective and less obvious way of detecting attempts to present oneself in the best possible way.**

The ? Scale: **Also known as "cannot say" scale, this validity scale is the number of questions left unanswered.**

6.4.2 The Minnesota Multiphasic Personality Inventory

The Trin scale: **this True Response Inconsistency Scale was developed to detect inconsistency.**

The VRIN Scale: **referred to as Variable Response Inconsistency Scale is another method to detect inconsistency.**

The Fb Scale: **this scale is composed of 40 items that less than 10% of respondents support. High scores on this scale sometimes indicate that the respondent stopped paying attention and began answering questions randomly.**

6.4.2 The Minnesota Multiphasic Personality Inventory

Supplemental Scales: **to supplement these multidimensional scales and to assist in interpreting the frequently seen diffuse elevations due to general factor designed to assess the extent to which a client admits to or is prone to abusing substances and A (anxiety) and R (repression) scales.**

PSY – 5 Scales: **These scales were developed as an attempt to connect with more general trend in personality psychology. The five components were labeled Negative Emotionality (NEGE), Psychoticism (PSYC), Introversion (INTR), Disconstraint (DISC) and Aggressiveness (AGGR).**

6.4.3 The Personality Adjective Checklist – PACL

The Personality Adjective Checklist is a comprehensive, objective measure of eight personality styles such as:

- **Introversive.**
- **Inhibited.**
- **Cooperative.**
- **Sociable.**
- **Confident.**
- **Forceful.**
- **Respectful.**
- **Sensitive**

This test is designed for normal adults, test reports are computer generated, and intended for use by qualified professionals only.

6.4.3 The Personality Adjective Checklist – PACL

PACL tests are considered probabilistic in nature and cannot be considered definitive. PACL personality scales measure theoretically derived, normal versions of the character types.

Test is exceptionally easy to administer and complete (5 – 10 minutes).

It is frequently used by personnel psychologists who work with relatively high functioning individuals and who want to understand strengths and weaknesses of their candidates.

6.4.4 The Cross Cultural Adaptability Inventory - CCAI

The CCAI is a culture general assessment designed to assess an individual's cross-cultural adaptability. The authors Colleen Kelley and Judith Meyers identified behavioral, emotional, and problem solving skills that are related to successful cross-cultural adaptation. It was originally created in 1987 and was revised in 1989 and 1992.

The CCAI is designed to respond to practical concerns expressed by culturally diverse and cross-culturally oriented people and by trainers who work with them. It is used as a part of a battery of interviews and testing.

6.4.4 The Cross Cultural Adaptability Inventory - CCAI

Scales

Emotional Resilience (ER) Scale **- helps measure the degree to which an individual can rebound and react positively to new experiences. It is the largest of the four, with 18 items.**

Flexibility / Openness (FO) Scale**: helps measure the extent to which a person enjoys the different ways of thinking and behaving that are typically encountered in the cross cultural experience. It has 17 items.**

6.4.4 The Cross Cultural Adaptability Inventory - CCAI

Scales

Perceptual Acuity (PAC) Scale **- helps measure the extent to which a person pays attention to and accurately perceives various aspects of the environment. It has 10 items deals with communication and perception cues across cultures.**

Personal Autonomy (PA) Scale**: helps measure the extent to which a person has evolved a personal system of values and beliefs and at the same time respect others and their value system. The scale has seven items.**

6.4.5 The NEO – PI -R
(Neuroticism, Extraversion and Openness Personality Inventory)

NEO PI-R is a psychological personality inventory, a 240 item measure of the five factor Model.

1. **Extraversion.**
2. **Agreeableness.**
3. **Conscientiousness.**
4. **Neuroticism.**
5. **Openness to experience**

The test was developed by Paul T Costa, Jr. and Robert R McCrae for use with adult (17+) men and women without overt psychopathology.

6.4.5 The NEO – PI -R
(Neuroticism, Extraversion and Openness Personality Inventory)

A list of the personality dimensions measured by NEO PI-R including facets is as follows

1. Neuroticism

- **Anxiety.**
- **Hostility.**
- **Depression.**
- **Self – Consciousness.**
- **Impulsiveness.**
- **Vulnerability to Stress.**

6.4.5 The NEO – PI -R
(Neuroticism, Extraversion and Openness Personality Inventory)

2. Extraversion

- **Warmth**
- **Gregariousness.**
- **Assertiveness.**
- **Activity.**
- **Excitement Seeking.**
- **Positive Emotion**

6.4.5 The NEO – PI -R
(Neuroticism, Extraversion and Openness Personality Inventory)

3. Openness to experience

- **Fantasy**
- **Aesthetics**
- **Feelings**
- **Actions**
- **Ideas**
- **Values**

6.4.5 The NEO – PI -R
(Neuroticism, Extraversion and Openness Personality Inventory)

4. Agreeableness

- **Trust**
- **Straightforwardness**
- **Altruism**
- **Compliance**
- **Modesty**
- **Tender mindedness**

6.4.5 The NEO – PI -R
(Neuroticism, Extraversion and Openness Personality Inventory)

4. Agreeableness

- **Fantasy**
- **Aesthetics**
- **Feelings**
- **Actions**
- **Ideas**
- **Values**

6.4.5 The NEO – PI -R
(Neuroticism, Extraversion and Openness Personality Inventory)

5. Conscientiousness

- **Competence**
- **Order**
- **Dutifulness**
- **Achievement Striving**
- **Self Discipline**
- **Deliberation**

6.4.5 The NEO – PI -R
(Neuroticism, Extraversion and Openness Personality Inventory)

Reliability

The internal consistency information of the test presented in the manual was derived from the full job performance sample (n = 1,539) Test retest reliability of the NEO PI-R is god. Costs and McCrae point out that this not only shows good reliability of the domains but also that they are stable over long periods of time (past the age of 30)as the scores over six years are only marginally more different than the scores as measured few months apart.

6.4.5 The NEO – PI -R
(Neuroticism, Extraversion and Openness Personality Inventory)

Validity

Numerous studies have evaluated the ability of NEO PI -R To predict important outcomes. Piedmont and Weinstein showed that (high) Neuroticism and low Conscientiousness were able to predict longitudinally levels of job burnout in a sample of 42 Occupational Therapists.

Miller demonstrated that levels of Neuroticism, Extraversion and Conscientiousness were significantly related to psychotherapy outcome. Levels of low Agreeableness have been linked longitudinally to coronary heart disease.

6.4.5 The NEO – PI -R
(Neuroticism, Extraversion and Openness Personality Inventory)

Norms

Normative information for Form S is based on a sample of 500 men and 500 women screened from a larger pool of 2,273 individuals. These 1000 individuals were selected demographically in order to match US Census projections for 1995.

Form R norms were obtained from 143 ratings of 73 men and 134 ratings of 69 women. These ratings were obtained from both spouses and multiple peer ratings (Costa & McCrae, 1992b)

6.4.5 The NEO – PI -R
(Neuroticism, Extraversion and Openness Personality Inventory)

Activity

– testing time:

Click on the link below for a free online version of the Revised NEO Personality Inventory

http://personality-testing.info/tests/NEO-PI-R.php

6.5 Interest and Motivation Test

6.5.1 The Strong Interest Inventory – SII

The Strong Interest Inventory – SII is an interest inventory used in career assessment. The test was developed in 1927 by psychologist EK Strong. Jr. to help people exiting military to find suitable jobs. It was revised later by Jo-Ida Hansen and David Campbell. The modern version is based on the typology of psychologist John L Holland.

It consists of 291 items. It is an assessment of interests. The test can typically be taken in 25 minutes after which the results must be scored by computer. After scoring, an individual can then view how their personal interests compare with the interests of people in a specific field.

6.5 Interest and Motivation Test

6.5.1 The Strong Interest Inventory – SII

The Strong Interest Inventory – SII is an interest inventory used in assessment of occupations.

The strong Interest inventory provides information from general interest types to specific occupations. The basic format of the Strong is to present various activities as below.

Activity	Like	Indifferent	Dislike
Visit an art museum	X		
Play golf		X	
Fix an auto engine			X
White water rafting		X	
Collect Stamps	X		

6.5 Interest and Motivation Test

6.5.1 The Strong Interest Inventory – SII

The responses would be collected into theme and activity scales based on content and a score for occupations would be derived by matching these responses to those of people in the various occupations.

Realistic Jobs	Investigative Jobs	Artistic Jobs	Social Jobs
Army Officer	Chemist	Artist	Nurse
Veterinarian	Pharmacist	Reporter	Teacher
Forester	Dentist	Interior Decorator	Minister

6.5 Interest and Motivation Test

6.5.2 The Kuder Preference Record – KPR

The Kuder Preference Record measures artistic, clerical, mechanical and scientific interests.

The Kuder Preference Record – KPR was one of the first interest inventories. The Kuder Assessments remain the widely accepted "standard" for adults to use when building their research-based educational and career plans.

It had 168 three choice items focusing on vocational interests which returned scores on ten scales which were claimed to measure such areas as artistic, clerical, mechanical and scientific interests.

6.5 Interest and Motivation Test

6.5.2 The Kuder Preference Record – KPR

The Kuder Occupational Interest Survey yields the person's scores along ten vocational interest scales.

6.5 Interest and Motivation Test

6.5.2 The Kuder Preference Record – KPR

The test results are presented as percentile scores and the report lists them separately for men and women. It then compares the person's scores on these scales to scores obtained by people holding certain professions and lists the top matches.

The test is a paper-and-pencil test that consists of 100 forced-choice triads of activities and takes about 30 minutes to complete.

MORE

http://www.kuder.com/default.aspx

6.6 Other Psychological Tests

6.6.1 The Rorschach Ink Blot Test

This is a psychological test in which subjects' perceptions of inkblots are used to test a person's personality characteristics and emotional functioning. The test is named after its creator, Swiss psychologist Hermann Rorschach. It is usually a battery of tests that often include the MMPI.

The Rorschach inkblot test is a type of projective psychological test used to analyze personality and emotional functioning

6.6 Other Psychological Tests

6.6.1 The Rorschach Ink Blot Test

Method

The tester and the candidates typically sit next to each other at a table, with the tester slightly behind the subject. There are ten official inkblots, each printed on a separate white card, approximately 18x24 cm in size.

Each of the blots has near perfect bilateral symmetry. Five inkblots are of black ink, two are of black and red ink and three are multicolored, on a white background. After the candidate has seen and responded to all of the inkblots (free association phase), the subject is shown them again.

6.6 Other Psychological Tests

6.6.1 The Rorschach Ink Blot Test

Method

The tester presents them one at a time in a set sequence for the subject to study. The subject is asked to note where he sees what he originally saw and what makes it look like that (inquiry stage).

The candidate is asked to hold the card and may rotate them. The psychologist writes down everything the candidate says or does. Analysis of responses is recorded by the test administrator using tabulation and scoring sheet.

6.6 Other Psychological Tests

6.6.1 The Rorschach Ink Blot Test

Method

The interpretation
a Rorschach record req[illegible] a wealth of knowledge
concerning personality
dynamics generally as well as
considerable experience with the
Rorschach method specifically.

6.6 Other Psychological Tests

6.6.1 The Rorschach Ink Blot Test

Exner Scoring System

Exner Scoring System also known as Rorschach Comprehensive System (RCS) is the standard method for interpreting the Rorschach test. It was developed in the 1960s by Dr. John E Exner, as a more rigorous system of analysis.

It has been extensively validated and shows high inter-rater reliability. The key components of the Exner system are the clusters of Rorschach variables and a sequential search strategy to determine the order in which to analyze them framed in the context od standardized administration, objective, reliable coding, and representative normative database.

6.6 Other Psychological Tests

6.6.1 The Rorschach Ink Blot Test

Validity

When interpreted as a projective test, results are poorly verifiable. The Exner system of scoring is meant to address this and has all but displaced many earlier (and less consistent) scoring systems.

Reliability

Depends substantially on testing procedure – where tester and the subject are seated, introductory words, responses to subject's queries etc. Exner has published detailed instructions.

6.6 Other Psychological Tests

6.6.1 The Rorschach Ink Blot Test

Activity – The Ink Blot Test

Click on the link below to take the test.

http://www.theinkblot.com

6.6 Other Psychological Tests

6.6.2 Thematic Appreciation Test – TAT

Thematic Appreciation Test – TAT is a projective psychological test. Historically, it has been among the most widely researched, taught and used of such tests. Its adherents assert that the TAT taps a subject's unconscious to reveal repressed aspects of personality, motives and needs for achievement, power and intimacy and problem solving abilities.

"Thematic Appreciation Test is a projective psychological test which taps a subject's unconscious to reveal repressed aspects of personality, motives and needs for achievement, power and intimacy, and problem solving abilities"

6.6 Other Psychological Tests

6.6.2 Thematic Appreciation Test – TAT

Procedure:

The TAT is popularly known as the picture interpretation technique because it uses a standard series of provocative yet ambiguous pictures about which the subject is asked to tell a story. The subject is asked to tell as dramatic story as they can for each picture presented; including

- **What has led up to the event shown**
- **What is happening at the moment**
- **What the characters are feeling and thinking**
- **What the outcome of the story was.**

6.6 Other Psychological Tests

6.6.2 Thematic Appreciation Test – TAT

Procedure:

What has led up to the event shown
What is happening at the moment
What the characters are feeling and thinking
What the outcome of the story was.

If the above elements are omitted, particularly for individuals of low cognitive abilities, the evaluator may ask the subject about them directly. The complete version of the test contains 31 picture cards. One card is completely blank.

Although the cards were originally designed to be matched to the subject in terms of age and gender, any card may be used with any subject.

6.6 Other Psychological Tests

6.6.2 Thematic Appreciation Test – TAT

Scoring Systems:

The TAT is a projective test in that, like the Rorschach test, its assessment of the subject is based on what he or she projects onto images which can be interpreted as the subject chooses. Therefore, to complete the assessment, each narrative created by a subject must be carefully recorded and analyzed to uncover underlying needs, attitudes, and patterns of reaction.

The test is used by the Services Selection Board of India.

6.6 Other Psychological Tests

6.6.2 Thematic Appreciation Test – TAT

Interpretation:

There are two basic approaches to interpreting responses to the TAT, called nomothetic and idiographic respectively. Nomothetic interpretation refers to the practice of establishing norms for answers from subjects in specific age, gender, racial or educational level groups and then measuring a given subject's responses against those norms.

Idiographic interpretation refers to evaluating unique features of the subject's view of the world and relationships.

6.6 Other Psychological Tests

6.6.2 Thematic Appreciation Test – TAT

Results:

The results of TAT must be interpreted in the context of the subject's personal history, age, sex, level of education, occupation, racial or ethnic identification, first language and other characteristics that may be important.

7. The Role of Industrial Psychologist

An important job of an industrial psychologist is to create testing rubrics for in-depth psychological testing that employees may take before getting an interview or the testing done by factories to see how quickly employees can recognize and decode tests and other information in a production setting.

Industrial-organizational psychology is the reasoning behind these tests and the people employed to create them do so to make sure that all employees are capable of doing the work required in that job setting.

8. Summary

- Psychological testing extends well beyond the use of intelligence and personality tests. Anything that requires a test taker to perform a behavior that is used to measure some personal attribute, trait or characteristic or to predict an outcome can be considered a psychological test.
- Psychological tests have various similarities and many differences. All psychological tests require an individual to perform one or more behaviors and these behaviors are used to measure some personal attribute, trait or characteristic which is important in understanding behavior or to predict an outcome.

8. Summary

- However, psychological tests can and do differ in terms of the behaviors they require individuals to perform, the attributes they measure, their content, how they are administered and formatted, how they are scored and interpreted and their psychometric quality.

- In organizational settings, human resource professional and industrial / organizational psychologists use psychological tests to make decisions such as whom to hire for a particular position, what training an individual needs and what performance rating an individual will receive.

8. Summary

- Psychological tests also played a key role in World War II that provided an impetus for developing of assessment centers. Since World War II, organizations have made use of psychological tests for selection and placement of new employees, evaluation of current employees and surveying customers.

- A test can be described as a standardized procedure for sampling behavior and describing it with categories or scores. In addition most tests have norms or standards by which the results can be used to predict other , more important behaviors.

8. Summary

- Tests always constitute a sample of behavior, never the totality of that which examiner seeks to measure. For this reason, test results always incorporate some degree of measurement error.

- Assessment is the process of compiling information about a person and using it to make inferences about characteristics and to predict behavior. Assessment incorporates testing but is more comprehensive and may include observations, interviews and other sources of information.

8. Summary

- Group tests are pencil-and-paper measures suitable to testing large groups of persons at one time. Individual tests are designed for one-on-one administration.; the examiner can thereby observe motivation and other characteristics of the candidate.

- Contrary to popular expectation, most studies find that the sex, experience and race of the examiner have little effect on psychological test results. Nonetheless, there may be specialized cases in which examiner-examinee interactions produce a detrimental effect on test scores.

8. Summary

- Standardized procedure is an essential feature of any psychological test. A test is considered to be standardized if the procedures for administering it are uniform from one examiner and setting to another.

- But concerns about the "fairness" of testing continues in many different settings. For every standardized test, there will be critics suggesting that the standardization prevents an illumination of the "essence" of the person. Others argue that lack of standardization permits favoritism.

8. Summary

- Psychological testing will always have a values component to it in addition to the issues related to content and process.

- Psychological testing can add to the productivity of an organization, particularly when incorporated as part of a will-structured human resource strategy aligned with business and organizational values, culture and objectives.

- **End of summary.**

9. Further Reading

‘The NFL Beat: Wonderlickin’ Bad

Wonderlic, Inc. administered its first exam in 1937. since then approximately 120 million prospective employees have taken this test across virtually every industry in existence. The test is used by employers as a general measure for gauging intelligence and problem solving aptitude present in potential members of their given workforce.

For more?

http://www.austinchronicle.com/blogs/sports/2012-04-05/the-nfl-beat-wonderlickin-bad/

Self Assessment Questions

1. In defining a test, the term objectivity refers to :

A Objective criteria used in scoring and interpretation.*

B An ideal rarely achieved.

C The degree of competence shown by the examiner.

D Uniformity of procedure.

2. Standardization means

A Objectivity of scoring.

B The use of standard scores.

C Experimental control over extraneous variables.

D Uniformity of procedure*

Self Assessment Questions

3. A battery consists of

A Tests used only for clinical assessment.

B Tests that are not standardized.

C Standardized tests like the Stanford-Binet.

D A group of tests.*

4. Major tests like the Stanford-Binet

A Are available for purchase by qualified users.*

B Are available for purchase by anyone.

C Are not proprietary tests.

D Are all criterion-referenced.

Self Assessment Questions

5. In considering ethical issues, three areas are most important

A Qualifications of the examiner, privacy and feedback.

B Informed consent, confidentiality and privacy.*

C Feedback, rapport and informed consent.

D Validity, reliability and norms

6. Administering a preliminary test form to a group of subjects

A Content analysis.

B Table of specifications.

C Standardization.

D Pilot testing.*

Self Assessment Questions

7. Factor analysis as a way of test construction assumes that

A Scales within an instrument correlate highly with each other.

B All items that make up a scale do not correlate with each other.

C Cross validation is not necessary.

D Scales should be univariate and independent.*

8. Item Response Theory assumes that

A A test score is the sum of true score and an error component.

B Responding to an item represents a reflex.

C Performance on a test is a function of an unobservable proficiency variable.*

D Both low scores and high scores can do equally well given the opportunity.

Self Assessment Questions

9. From the point of test construction, which of these does NOT belong?

A Myers-Briggs Type Indicator.

B Edwards Personal Preference Schedule

C California Psychological Inventory.*

D Personality Research Form.

10. The letters NEO stand for

A Normality, Exceptionality, Orderliness.

B The three authors (Norman, Everett and Oserinsky.)

C New England Observational (scale).

D Neuroticism, Extraversion, Openness to experience.*

Self Assessment Questions

11. The 15 subtests of the 1986 Stanford-Binet are administered

A With the Vocabulary test last.

B In a specific sequence.*

C With short term memory test first.

D As examiner deems appropriate.

12. A set of six scales designed to portray a "general" type (like realistic or artistic) on the strong

A Basic Interest scales.

B Occupational scales.

C Administrative scales.

D General Occupational themes.*

Self Assessment Questions

13. The Kuder was originally developed on the basis of

A Criterion group scaling.

B Homogenous scaling.*

C Criterion-keying.

D Longitudinal studies.

14. The Lie scale on the MMPI is composed of items that

A Have high correlation with social desirability.

B Are quite heterogeneous but cover poor physical health.

C Showed a significant response shift between lie and honest instructions.

D Most people, if answering honestly, would not endorse.*

Self Assessment Questions

15. The Wonderlic personnel Test is designed to assess

A How well organized and dependable the person is.

B General mental ability.*

C Motivation for advancement.

D Leadership ability.

Thank You !

S.P MANDALI'S
weschool
Welingkar Education