

# SAMPLE

## INTRODUCTION

---

### Aim

Explain the main kinds of psychological tests and why they are used, and meaning of test reliability and validity.

Psychological tests assess and evaluate information that is given to the examiner by the client. This information is in the form of answers to interview questions or as answers on paper—or on a computer—to specific questions. Ultimately, a test's accuracy depends on how carefully and seriously the client answers the questions they are asked.

Psychological tests are changing all of the time. New tests are developed. Existing tests are updated. It is important to stay abreast of new tests in the area in which you plan to work.

### WHAT IS A PSYCHOLOGICAL ASSESSMENT?

A psychological assessment is when a skilled psychologist uses the techniques and tools of psychology to learn general or specific facts about another person, either to inform others how they function or to predict their behaviour and functioning in the future.

Mahoney and Ward stated that assessments –

- Often use tests
- Typically, do not follow defined procedures or steps
- Contribute to a decision making process to some problem, often be redefining a problem, breaking the problem down into smaller pieces or highlighting parts of a problem
- Require that the examiner considers, evaluates and integrates the data
- Are less inflexible and routine
- Are individualised
- Produce results that cannot be solely evaluated on psychometric grounds

Assessment is usually for the point of diagnosis or classification. That is, placing a person into a loosely or strictly defined category of people. This allows us to understand what they are like in general and assess the presence of other relevant characteristics based on people similar to them.

### TYPES OF PSYCHOLOGICAL TESTS

The understanding behind psychometric tests is that if things such as intelligence and personality exist, we should be able to measure them. Difficulties with this will be discussed later.

Psychological tests fall into several categories:

**Achievement tests** are usually seen in educational or employment settings, and they attempt to measure how much you know about a certain topic (i.e., your achieved knowledge), such as mathematics or spelling.

**Aptitude tests** are also generally used in educational and employment settings, and they attempt to measure how much of a capacity you have (i.e., your aptitude) to master material in a particular area, such as mechanical relationships.

**Intelligence tests** attempt to measure your intelligence, or your basic ability to understand the world around you, assimilate its functioning, and apply this knowledge to enhance the quality of your life. Intelligence, therefore, is

a measure of a potential, not a measure of what you've learned (as in an achievement test), and so it is supposed to be independent of culture. The trick is to design a test that can actually be culture-free; most intelligence tests fail in this area to some extent for one reason or another.

**Neuropsychological tests** attempt to measure deficits in cognitive functioning (i.e., your ability to think, speak, reason, etc.) that may result from some sort of brain damage, such as a stroke or a brain injury.

**Occupational tests** attempt to match your interests with the interests of persons in known professions. The logic here is that if the things that interest you in life match up with, say, the things that interest most school teachers, then you might make a good school teacher yourself.

**Personality tests** attempt to measure your basic personality style and are most used in research forensic settings to help with clinical diagnoses. Two of the most well-known personality tests are -

- the Minnesota Multiphasic Personality Inventory (MMPI), composed of several hundred “yes or no” questions, and
- the Rorschach (the “inkblot test”), composed of several cards of inkblots—you simply give a description of the images and feelings you experience in looking at the blots.

Specific clinical tests attempt to measure specific clinical matters, such as your current level of anxiety or depression.

There are many different psychological assessment tests available – below are some examples.

### **Aptitude/Intelligence/Educational Tests**

- Carl Hollow Square Test
- Cattell Culture Fair Intelligence Tests
- Cognitive Abilities Test 3 (CAT3)
- Collins-Drever Test
- Colour Trails Test (CTT)
- Communication and Symbolic Behaviour Scales (CSBS)
- Daniels - Figure Reasoning Tests
- Differential Aptitude Tests
- Draw a Person Test (DAP)
- Edinburgh Handedness Inventory
- Edinburgh Semantic Differential Scales
- Ekstrom Factor Referenced Cognitive Tests
- Elithorn Perceptual Maze Test
- Raven’s Standard Progressive Matrices Plus (SPM Plus)
- Renfrew Word-finding Vocabulary Scale. 3rd ed.
- Rey Auditory and Verbal Learning Test
- Reynell Development Language Scales
- Reynell Developmental Language Scales III (RDLS)
- Seashore Measures of Musical Talents
- The Speed and Capacity of Language-Processing Test (SCOLP)
- Stanford-Binet Intelligence Scale

## Personality Tests (Including occupational)

- Allport, Vernon & Lindzey Study of Values
- Bass - Famous Sayings (FS) Kit
- Bell Adjustment Inventory
- Bem Sex-role Inventory
- Brook Reaction Test of Interests & Temperament
- California Psychological Inventory (CPI)
- Cattell 16 Personality Factor Questionnaire 1979 (16PF)
- Cattell 16 Personality Factor Questionnaire 5th ed (16PF-5)
- Wilson-Patterson Attitude Inventory (WPAI)
- Work Aspect Preference Scale
- Zuckerman Sensation Seeking Scale (SSS-V)

## Memory Tests

- Autobiographical Memory Interview (AMI)
- Benton Test of Three Dimensional Constructional Praxis
- Benton Visual Retention Test
- Coping with Memory Problems : a practical guide for people with memory impairments, their relatives, friends and carers
- Working Memory Test Battery for Children (WMTB-C)
- 
- 

### New Research On Memory

Just a note here on memory. New research in Japan and America potentially may turn current theories of memory on its head. A team at Riken-MIT Center for Neural Circuit have done research and found that the brain actually makes two memories of events. One that is for the here and now and the other that will last the person's life time.

In the past it has been thought that our memories start in the short term memory and then are converted into the long term memory. However, this research suggests that this might not be the case.

In the past, it was thought that the hippocampus is where we store short term memories whilst the cortex is where we store long term memories. However, this new research has found that this is not the case. The research was carried out on mice and is in the very early stages of development, but if this research is correct, it could change how we think about how our memories work. It will be interesting to find out how this impacts upon psychological assessment and tests that look at memory in the future.

## Mood (Non-clinical)

- Berkeley Expressivity Questionnaire
- Emotion Regulation Questionnaire (ERQ)
- Profile of Mood States (POMS)
- Spielberger State-Trait Anger Expression Inventory
- Spielberger State-Trait Anger Expression Inventory-2
- Spielberger State-Trait Anxiety Inventory
- UWIST Mood Adjective Checklist

## Health Psychology and Clinical Scales

- Autism Diagnostic Interview Revised ADI-R (for use only by those who have completed training to use this test)
- Behavioural Assessment of the Dysexecutive Syndrome (BADS)
- Bender Visual Motor Gestalt Test
- Bene-Anthony Family Relations Test
- Broadbent Cognitive Failures Test
- Caine & Hope Hysteroid-Obsessoid Questionnaire (HOQ)
- Caine & Foulds Personality Questionnaire (HDHQ)
- Clinical Explanation of Language Functions - Preschool (CELF)
- Delis-Kaplan Executive Function System (D-KEFS)
- Dutch Eating Behaviour Questionnaire
- Eating Disorder Inventory-2 (EDI-2)
- Family Environment Scale (FES)
- Farnsworth-Munsell 100 Hue Test
- Stirling Eating Disorder Scales (SEDS)
- Stroop Neuropsychological Screening Test
- Structured Clinical Interview of DSM (Diagnostic and Statistical Manual) (SCID)
- Symbol Digit Modalities Test
- Telephone Interview for Cognitive Status (TICS)
- Vineland Adaptive Behaviour Scales (Vineland-II)
- The Visual Object and Space Perception Battery (VOSP)

These are just a summary of some tests. New tests are being developed all of the time or existing tests are being updated.

## DSM

DSM is the Diagnostic and Statistical Manual. It is produced by the American Psychiatric Association and provides a common language and criteria for mental health disorders throughout the world. But please note it is not the only diagnostic manual used. The DSM is regularly updated, so if you are carrying out research using the DSM, please ensure that you use the most regular version.

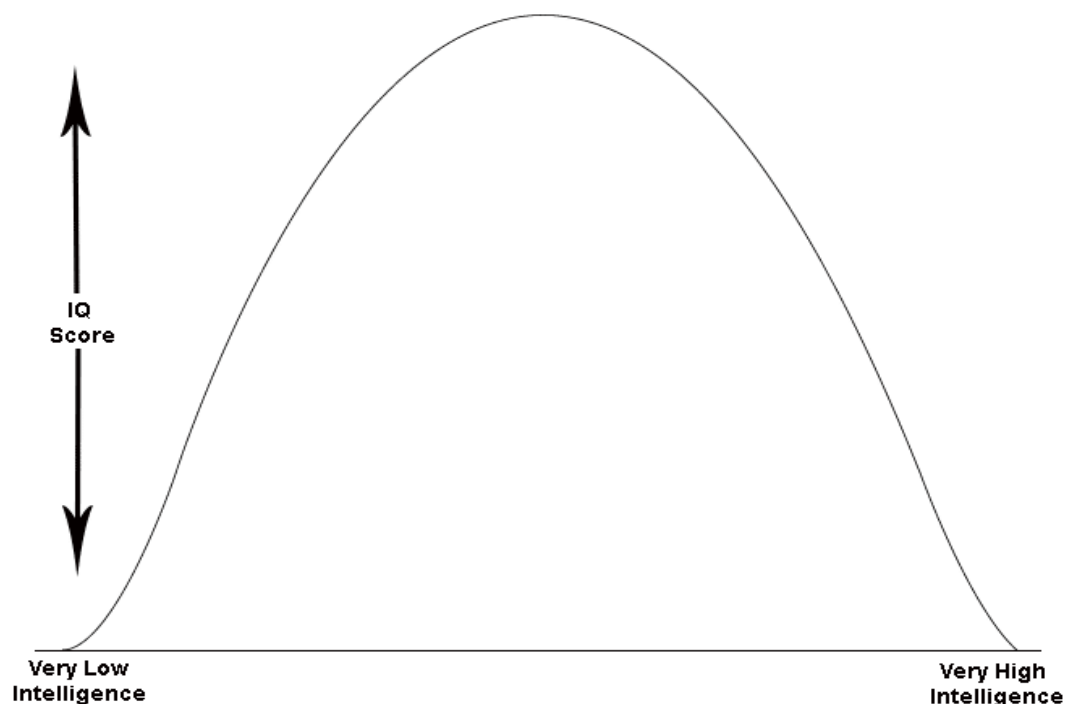
## History of Psychological Testing

Intelligence tests have been widely used psychological tests. It is interesting to consider the history of the development of intelligence tests.

The idea that intelligence should be or could be tested began with Sir **Francis Galton**, a nineteenth century British scientist. Galton was interested in many different fields including biology and psychology. After the publishing of Darwin's book "The Origin of the Species", Galton tried to discover a relationship between human ability and heredity.

At the time, it was thought that there were a small number of geniuses and a small number of people with very low intelligence, but that the majority were equally intelligent people. Whatever a person achieved in life was thought to be due to willpower and hard work. However, Galton was not satisfied with this as he believed that mental traits were based on physical factors and were inheritable traits.

Galton was influenced by the work of a Belgian statistician named Lambert Adolphe Jaques Quetelet, who was the first to apply statistical methods to human characteristics and discovered the concept of normal distribution. This is the tendency for the bulk of the population to fall between two extremes, with numbers dropping at either extreme. This is known as a bell curve.



Therefore, if you took the IQ of everyone in the world, the top 2% (approximately) would be very intelligent, the bottom 2% would have a very low IQ, whilst everyone else would fall between these two extremes, so with lower, some with higher IQs.

Galton wrote a book called “Hereditary Genius” which is recognized as the first scientific investigation into intelligence. In 1890s, an American student of Galton’s, called [Cattell](#), took the idea of intelligence testing to America. His work caused brief, but intense mental testing in America. However, the downfall of the test was that scoring well on the test did not indicate that a student would do well on schoolwork, which was considered practical proof of good mental ability.

Around the same time, [Alfred Binet](#) devised tests to measure child intelligence. His understanding of intelligence was developed through trial and error testing of local students. He worked with students with average intelligence and learning disabilities/mental handicap. He discovered that certain tasks that average students could handle, could not be done by the students with learning disabilities/mental handicap. Binet calculated the normal abilities for students at each age and could pinpoint their mental age by comparing it with the norm.

The educational authorities in Paris asked Binet to devise a test to enable them to separate children with “normal” educational abilities and those with special needs. His idea that the test could determine a child’s mental age was very popular. **Wilhelm Stern**, developed a better way of expressing results than mental age, based on the subjects chronological age and their mental age. For example, a ten year scoring one year ahead would not be as advanced as a five year old scoring one year ahead.

Lewis Terman developed the term **intelligence quotient (IQ)**. An average IQ score on Binet's test was 100. To score above 100 is seen as above average, whilst below 100 is below average. Binet's test had its limitations, so Binet and Stern doubted that IQ scoring would actually represent the innate quantity of intelligence. Stern wrote in 1914 -

"No series of tests, however skilfully selected it may be, does reach the innate intellectual endowment, stripped of all complications, but rather this endowment in conjunction with all influences to which the examinee has been subjected up to the moment of testing."

However, the Binet test was still accepted in America. In 1916, the test was administered to a prisoner on trial for murder. As the prisoner performed so badly, the jury acquitted him because of his mental condition.

In 1917, the America IQ testing grew when America entered World War I. Binet's tests were originally designed for children, but the army needed to sort out huge numbers of draftees into various Army positions. To solve this problem, they devised a mass intelligence test using seven leading psychologists. The Chairman was **Robert Yerkes**. Lewis Terman was also one of the seven. Terman had a pupil named **Arthur Otis** who had begun to devise a group intelligence test when the Army decided to begin to use one. The committee mainly adopted Otis's material. By 1919, two million American men had been tested. The Army scores were given marks for correct answers and divided into five classes, A to E.

Soon after World War I, many companies began using the programs to see who should be hired, promoted or transferred. The greatest market for intelligence tests was schools.

Walter Lippman, however, criticised the tests. He said -

*"One only has to read around in the literature of the subject...to see how easily the intelligence test can be turned into an engine of cruelty, how...it could turn into a method of stamping a permanent sense of inferiority upon the soul of a child..."*

The IQ began to fall out of favour in the 1960s and 1970s as it was racially and culturally specific in terms of questions. However, the concept of intelligence has continued to grow and evolve, despite the problems with and misuse of IQ tests. **Howard Gardener** has identified seven distinct forms of intelligence –

- logical-mathematical
- linguistic
- spatial
- musical
- bodily kinaesthetic
- interpersonal
- intrapersonal

This idea of multiple intelligences has moved intelligence away from a mathematical and verbal understanding. Gardener's ideas are used in many areas of education, but traditional IQ tests are still used in many areas.

## JUSTIFICATION FOR USING TESTS

Psychological tests were created for three main reasons, all of which are interconnected:

- It is easier to get information from tests than by clinical interview.
- The information from tests is more scientifically consistent than the information from a clinical interview.
- It is harder to get away with lying on a test than in a clinical interview.

## Advantages and Disadvantages of Using Psychometric Assessments

**Advantages** – Psychometric tests provide information on the performance of certain groups of people perform. They tend to be standardised tests. The test score of an individual is then compared with the test scores of a group of people. For example, we could test a hundred people and find that the average IQ was 100. If the individual was then tested and their score was 94, we would say that they were below average compared to the average results.

In the tests, all the people completing the test have the same instructions and the same task to do. This makes sure that any differences in performance are due to the person completing the task, not the task itself. To explain this more fully, if an interviewer sat and asked ten people the same question, there may be different results.

For example, the interviewer may not like the first respondent, so might rush through the test, or might explain things more fully to another respondent, or help another respondent. Whereas if all the respondents have the same instructions and task, the differences should be due to them, not because of the way they were asked the questions. Detailed instructions are also given to the people who administer and score the tests, so there should be some standardisation.

The tests are useful, because in theory, they allow us to make statistically based, objective judgements, on things such as, aptitude for jobs.

However, there are some difficulties with psychometric tests –

**Disadvantages** - As already mentioned, if things such as intelligence and personality exist, we should be able to measure them. The difficulty is, whether these things do actually exist, and if they do, whether it is possible to measure them. After all, IQ was invented by psychologists, so psychologists are then investigating something they consider important.

The person designing the test decides what should be in the test, so the test only measures what the designer thinks is important. If there is no agreed-upon definition, e.g. what would make a good teacher, and then different test designers may include different types of question.

Are personality traits stable? For example, we may perform really well in an IQ test one day, but perform poorly the next. Basically, psychological tests measure our performance in a certain task on a certain day.

Also, consider that values can be attached to judgements. Going back to the IQ scores above. The top 2% of the population may have a very high IQ, whilst the bottom 2% will have a very low IQ. However, having a very high IQ is considered to be a positive thing, whilst a very low IQ is considered a negative thing. Therefore, scores do not just give a figure for an IQ score, but have with it an attendant value.

## Ethnicity and Different Cultures

Some tests are not balanced for age and ethnicity, so they may not be fair. For example, a child may not be able to answer the same questions as an adult. Most tests now do try to take account of these different factors.

Finally, some tests are culturally biased. For example, a person living in America may never have seen the ocean in real life. Therefore, they may be less able to answer questions on the ocean than, say, a person who lives in the UK. This would create a bias; the American respondent may get a question wrong simply because a view of the ocean is outside their experience

The American Psychological Association offers the following brochure for details on why and how children with a different language or culture should be tested:-



## Psychological Testing of Language Minority and Culturally Different Children

This brochure provides information on psychological tests, generally, and on the importance of ensuring that tests are appropriate with regard to a child's cultural background and language, specifically. Also included are suggestions which would help parents to assure the use of appropriate psychological tests.

### What are psychological tests?

Psychological tests allow psychologists to assess your child's abilities, behaviour, and overall functioning. Tests compare how your child is developing with how other children of the same age are developing.

Testing can provide helpful and necessary information to plan for your child's developmental and educational needs. Testing can also help your child's teachers and others identify areas that need special attention, as well as areas of great potential.

### Why should a child be tested?

Parents, teachers, and others can have many questions about a child, including questions about behaviour, feelings, motivation, ability to learn, intellectual potential, brain functioning, interests, and problems. Specific tests have been developed to find answers to the many questions that you or others might have about your child.

### Can all tests be applied to all children equally?

No. Special consideration should be observed when testing language minority and culturally different children.

### What is a language minority child?

A language minority child is one whose first language is not English and who is not fluent in English.

### What is a culturally different child?

A culturally different child is one who lives in a home where the customs and values may be markedly different from the majority of the children in the social group used by the test as the basis for the comparison. Such children may even have been raised in another county.

### Why are some tests not right for a child whose language or culture is different?

Tests give information in the form of comparisons with other children of the same age. If your child's language and culture differ from those of the children he or she is being compared with, the comparisons may suggest conclusions that are not accurate. For example, if your child does not understand English well, test scores may not be correct. The examiner might incorrectly conclude your child has a problem in the area measured by the test.

Test items may also relate to experiences that are unfamiliar to a culturally different child. For example, in some cultures children are taught to respond to adults with short answers. If a psychological test requires this culturally different child to give long answers, then he or she may appear not to be doing well.

### Why it is important for parents to know about testing?

It is important that you help the examiner to make the best possible assessment of your child. You can do so by raising your concerns and questions about your child. Work with the psychologist to learn what the psychological tests are intended to measure.

Best possible measures of your child are obtained by the use of tests that (a) get the specific questions raised by you, the teacher, therapist, and others and (b) compare your child appropriately.

Use of tests or the interpretation of test results without consideration of language and culture might mean that your child is placed in an inappropriate school program or does not get the psychological treatment needed.

### *What can you do to assure that appropriate use of psychological tests?*

First, talk with the psychologist evaluating your child:

1. Ask if the tests being used are being given in the language that your child understands best.
2. Ask if the tests take into account your child's cultural background.
3. Ask if the tests used can compare your child to other children from the same cultural group.
4. Ask the psychologist working with your child if he or she has training and experience with your child's language and culture.
5. Ask if other information besides test scores will be used to make decisions about your child.

### **What if you have questions or concerns?**

If you have concerns about the response to these questions, the evaluation of your child may need to be reviewed. Here are some steps that you can take:

- You can request that your child's evaluation be postponed until it can be done appropriately.
- You can ask to speak to the psychologist's supervisor about this matter.
- You can express your concern in writing and request the psychologist's response.
- You can seek further guidance and assistance by contacting:

The American Psychological Association  
Committee on Psychological Tests and Assessment  
750 First Street, NE  
Washington, DC 20002-4242  
(202) 336-5500

### **VALIDITY AND RELIABILITY**

The main problem with psychological tests is whether they measure what they are supposed to measure. For this purpose we look at the validity and reliability of the tests.

#### **Validity**

Validity is a measure to determine how well the test measures what it is trying to measure. Psychological assessments are an important part of treatment and research, so it is important that the test measures what we want it to measure or we think it measures. For example, what if we wanted a test that looks at the stable and permanent emotions a person feels, but instead the test measures the emotions they are only feeling at that moment. Validity isn't measured by one statistic, it is measured by research that researches the test and the behaviour it is intended to measure. There are different types of validity -

**Construct Validity** is the most important kind of validity. It refers to the ability of the test to measure the psychological construct that it is designed to measure, for example, depression. We can assess this through the convergent or divergent validity of the test. This refers to whether a test can give results similar to other tests of the same construct and different from tests of different constructs. Establishing whether a test has construct validity is a long, complex process. The various qualities that contribute to construct validity include –

- Convergent validity
- Criterion validity
- Discriminant validity

To create a measure of construct validity, first the domain of interest must be defined, that is, what is to be measured, such as depression. Then we construct a measurement item that is designed to measure the domain.

Then there is a process of rigorous scientific testing and modifying the measure. In psychological testing, there is a bias towards selecting items that can be written down objectively.

**Convergent Validity** This considers whether tests will return similar results to other tests, which aim to measure the same or related constructs. So we consider does the measure match an external criterion e.g. Behaviour or another established test? Does it predict this “behaviour”?

**Divergent Validity** is the opposite of criterion validity. If you are measuring a construct believed to have no relationship to something else. If there is no relationship and if you measure has a good construct validity, you would expect scores to be unrelated to scores on a measure for a divergent construct.

**Criterion-related Validity** refers to the ability of a test to predict someone’s performance on something. For example, before actually using a test to predict whether someone will be successful at a particular job, you would first want to determine whether persons already doing well at that job (the criterion measure) also tend to score high on your proposed test. If so, then you know that the test scores are related to the criterion. Criterion validity includes concurrent and predictive validity.

**Concurrent validity** considers whether a measure relates to other manifestations of the construct that the test is supposed to measure.

**Predictive validity** asks whether the test predicts an individual’s performance in specific abilities.

**Content Validity** refers to the ability of a test to sample adequately the broad range of elements that compose a particular construct.

## Reliability

In psychology generally, reliability refers to how consistent a measure is. All measurement procedures have a potential for error, so the aim is to minimise this. An observed test score is therefore made up of a true score and a measurement error. The goal of estimating consistency (reliability) is to determine how much variability in the test is due to measurement error and how much is due to variability in the true score. Measurement errors are usually random, a person’s test score may not reflect their true score because they are ill, hungover, anxious, the room is noisy, they are distracted etc.

A test is considered to be reliable if we get the same result repeatedly. For example, if we have a test to measure a trait such as extraversion, then every time we administer a test to a subject, the results should be approximately the same. Therefore, it is impossible to calculate reliability exactly, but there are ways to estimate reliability.

The ability of a test to give consistent results is therefore known as its reliability.

Reliability can be improved by –

- Getting a range of different measures using slightly different methods and techniques.
- Getting repeated measures using the same test.

**Internal Consistency Reliability** refers to how well all the test items relate to each other.

**Test-retest Reliability** refers to how well results from one administration of the test relate to results from another administration of the same test at a later time.

Note that **without reliability, there can be no validity.**

Note also that no psychological test is ever completely valid or reliable because the human psyche is just too complicated to know anything about it with full confidence. That’s why there can be such uncertainty about a case even after extensive testing.

Also, if a client keeps “failing” a psychological test, for example, to get a particular job, it may be that they are not suited for that particular job, or it could be that they require psychological counselling to remove any underlying issues in their psyche.

### **SET TASK 1**

Carry out a library or internet search on psychological tests or psychological assessments. Try to find the most up to date tests you can find. Make a list of at least 10 new tests.

### **SET TASK 2**

Carry out a library or internet search on guidelines for psychological testing or assessment. For example, the British Psychological Society offer guidelines on testing and test use. Make notes.

### **SET TASK 3**

We mentioned the Rorschach ink blot test in this lesson. Find out more about the test online or in the library.