# Quadrant II – Transcript and Related Materials

**Programme: Bachelor of Arts (Third Year)**

**Subject: Psychology**

**Paper Code: PSC 106**

**Paper Title: Psychological Testing**

**Unit: Test Development**

**Module Name: Test Construction and Tryouts.**

**Name of the Presenter: Maria Muriel de Barros (Assistant Professor)**

**Rosary College of Commerce and Arts, Navelim-Goa**

**Notes**

## TEST CONSTRUCTION AND TEST TRYOUTS

Writing items for computer administration:

A number of widely available computer programs are designed to facilitate the construction of tests as well as their administration, scoring, and interpretation. An item bank is a relatively large and easily accessible collection of test questions. Instructors who regularly teach a particular course sometimes create their own item bank of questions that they have found to be useful on examinations.

One of the many potential advantages of an item bank is accessibility to a large number of test items conveniently classified by subject area, item statistics, or other variables. Items may be added to, withdrawn from, and even modified in an item bank

Computerized Adaptive Testing (CAT): interactive computer-administered test-taking process wherein items presented to the test taker are based in part on

the test taker's performance on the previous items. In Computerized Adaptive Testing (CAT) only a sample of the total number of items in the item pool is administered to any one test taker, based on their previous response patterns. Items that have a high probability of being answered in a particular fashion ("correctly" if an ability test) are not presented, thus providing economy in terms of testing time and total number of items presented.

CAT tends to reduce *floor effects* and *ceiling effects.* A *floor effect* refers to the diminished utility of an assessment tool for distinguishing test takers at the low end of the ability or trait being measured. A *ceiling effect* refers to the diminished utility of an assessment tool for distinguishing test takers at the high end of the ability, trait being measured.

The ability of the computer to tailor the content and order of presentation of test items on the basis of responses to previous items is referred to as *item branching*. A computer that has stored a bank of achievement test items of different difficulty levels can be programmed to present items according to an algorithm or rule.

## C) SCORING ITEMS

Scoring items is the third aspect of test construction. There are several methods of allocating scores to test items.

1. Cumulative model: The higher the score on the test, the higher the test taker is on the ability, trait, or other characteristic that the test purports to measure.
2. Class or category scoring: Test taker response earns credit towards placement in a particular class or category with other test takers whose pattern of responses is presumably similar in some way.

- For example: The class or categorical scoring approach is used by some diagnostic systems wherein individuals must exhibit a certain number of symptoms to qualify for a specific diagnosis.

3. Ipsative scoring*:* A typical objective in is comparing a test taker's score on one scale within a test to another scale within that same test. To draw intra-individual conclusions about the test taker. For example : "John's need for achievement is higher than his need for affiliation."

**TEST TRYOUTS**

The third stage of test development is test tryouts. In this stage the final version of the test is developed from a pool of items and the test developer will try out the test on a representative sample. The test should be tried out on people who are similar in critical respects to the people for whom the test was designed.

For example: A test designed to aid decisions regarding the selection of corporate employees with management potential at a certain level, it would be appropriate to try out the test on corporate employees at the targeted level.

The number of subjects for try outs: An informal rule of thumb is that there should be no fewer than five subjects and preferably as many as ten for each item on the test. The more subjects in the try out the better as the role of chance is weaker in the subsequent data analysis.

The try-out must be executed under conditions as identical as possible to the conditions under which the standardized test will be administered; (instructions, time limits allotted, the atmosphere at the test site)

**A GOOD TEST ITEM:** A good test is reliable and valid. It helps to discriminate between test takers. A good test item is one that is answered correctly by high

scorers and incorrectly by low scorers on the test as a whole. Test developer analyses test scores and responses to individual items.

The different types of statistical scrutiny that the test data can potentially undergo are referred to collectively as item analysis, which is the next step in test development