

Content Validity

- Test blueprint
 - A plan regarding the types of information to be covered by the items, the number of items tapping each area of coverage, the organization of the items in the test.
 - also called Table of Specifications (TOS)
 - **Employment test** to be content-valid, its content must be a representative sample of the job-related skills required for employment.
 - **Behavioral observation** is one technique frequently used in blueprinting the content areas to be covered in certain types of employment tests.

TABLE OF SPECIFICATIONS
TEST IN THEORIES OF PERSONALITY
Midterm Examination
1st Semester SY 2019-20

Topics		Level of Objectives (No. of Items)			Item location	Total no of Items	Percent
	Specific Learning Outcomes	Comprehension	Analysis	Evaluation			
Introduction	<u>Discuss</u> the role of personality theory in psychology.	7	6	6	1-19	19	27.14
Sigmund Freud: Psychoanalysis	<u>Explain</u> concepts and process in Psychoanalysis	14	10	5	20-48	29	41.43
Carl Jung: Analytical Psychology	<u>Discuss</u> Jung's view of personality and <u>differentiate</u> it from Freud's view.	4	2		54-57 67-68	6	8.60

Criterion-Related Validity

- **What Is a Criterion?**

- Criterion as the **standard** against which a test or a test score is evaluated.
 - Example, if a test purports to measure the trait of *athleticism*, we might expect to employ “*membership in a health club*” or any generally accepted measure of physical fitness
- There are ***no hard-and-fast rules*** for what constitutes a criterion.
 - It can be a test score, a specific behavior or group of behaviors, an amount of time, a rating, a psychiatric diagnosis, a training cost, an index of absenteeism, an index of alcohol intoxication, and so on.

Decision theory and Test Utility

- **Base rate**
 - is the extent to which a particular trait, behavior, characteristic, or attribute exists in the population (expressed as a proportion).
 - Due consideration must be given to the base rate of a targeted attribute in the sample of people being studied in predictive validity research as compared to the base rate of that same attribute in the population at large.
- **Hit rate** may be defined as the **proportion of people** a test **accurately identifies as possessing or exhibiting** a particular trait, behavior, characteristic, or attribute.
 - For example, hit rate could refer to the proportion of people accurately predicted to be able to perform work at the graduate school level or to the proportion of neurological patients accurately identified as having a brain tumor.
- **Miss rate** may be defined as the proportion of people the test **fails to identify** as having, or not having, a particular characteristic or attribute.
 - amounts to an **inaccurate prediction**.

Construct Validity

- Evidence of change with Age
- Some constructs lend themselves more readily to predictions of change over time.
 - Giftedness or intelligence
 - Reading ability
 - Marital satisfaction
- Measures of marital satisfaction may be less stable over time or more vulnerable to situational events than is reading ability.
- Evidence of change over time, like evidence of test homogeneity, does not in itself provide information about how the construct relates to other constructs.

Construct Validity

- Evidence from distinct groups
 - Also referred to as **Method of Contrasted Groups**
 - One way of providing evidence for the validity of a test is to demonstrate that scores on the test vary in a predictable way as a function of membership in some group.
 - The rationale here is that ***if a test is a valid measure of a particular construct, then test scores from groups of people who would be presumed to differ with respect to that construct should have correspondingly different test scores.***
 - *Consider a test of depression. We would expect individuals psychiatrically hospitalized for depression to score higher on this measure than a random sample of people.*

Validity, Bias, and Fairness

- **Rating error**
 - A **rating** is a numerical or verbal judgment (or both) that places a person or an attribute along a continuum identified by a scale of numerical or word descriptors known as a rating scale.
 - A **rating error** is a judgment resulting from the intentional or unintentional misuse of a rating scale.
 - **Leniency error** (also generosity error)
 - an error in rating that arises from the tendency on the part of the rater to be lenient in scoring, marking, and/or grading.
 - **Severity error**
 - Less than accurate rating or error in evaluation due to the rater's tendency to be overly critical.