PSY443- TESTING AND ASSESSMENT IN PSYCHOLOGY

VALIDITY

Validity & Reliability

- ✓ Both are properties of a measure
- ✓ ‘How good is this measure?’

First approximation:

- ✓ Reliability = **Dependability**
  - ✓ Validity = **Truthfulness or Accuracy**
- ✓ Measuring the accuracy and consistency of your research instruments - known as validity and reliability

VALIDITY

VERSUS

RELIABILITY

VALIDITY	RELIABILITY
Extent to which a test measures, and what it claims to measure	Consistency of the test results
Measures whether the extent to which the test claims to measure is achieved	Measures whether the test gives consistent results at its outcome
There are two types as internal validity and external validity	There are two types as internal reliability and external reliability

VALIDITY

- To measure what it is designed to measure.
- **A VALID TEST** measures - what it is intended to measure.

Why Validity?

Validity is done mainly to answer the following questions:

- ✓ Is the research investigation providing answers to the research questions for which it was undertaken?
- ✓ If so, it providing these answers using appropriate methods and procedures?

Bases of Validity in Quantitative Research

- Controllability
- Replicability
- Predictability
- Generalizability
- Randomization of sample
- Neutrality
- Objectivity
- Observability
- Inference
- Manipulation of variables

Types of Validity

1. External validity
2. Internal validity
3. Content validity
4. Criterion validity
5. Construct validity
6. Population validity
7. Ecological validity

Content Validity (Kapsam Geçerliği)

- The measurement method covers the entire range of relevant behaviors, thoughts, and feelings that define the construct being measured.
  - Example: One's attitude toward an object is considered to consist of thoughts about the object, feelings about the object, and behaviors toward the object.

- A test to assess one's attitude toward taxes should include items about thoughts, feelings, and behaviors.

- If test anxiety is thought to include both nervous feelings and negative thoughts, then any measure of test anxiety **should cover** both of these aspects.

Construct Validity (Yapı Geçerliği)

- The measurement tool operates correctly on a conceptual level.

Subtypes:

- **Convergent Validity**: This subtype examines whether a test correlates strongly with other measures that theoretically should be related to it.
  - For example, if you're measuring self-esteem, the test should have high correlations with other validated self-esteem measures. High convergent validity indicates that the tool is accurately capturing the intended construct.
- **Divergent (or Discriminant) Validity**: Divergent validity tests whether the measurement tool is not strongly correlated with measures of different, unrelated constructs.
  - For instance, a self-esteem measure should not correlate highly with a measure of intelligence, as they are distinct constructs.
- **Factorial Validity**: Factorial validity is established through factor analysis, a statistical method used to see if the test items group into the theoretical dimensions or factors that the construct is thought to have.
  - For example, a measure of mental health might be theorized to have factors like anxiety and depression, and factor analysis should reveal these distinct clusters if the construct validity is strong.

Criterion Validity

- It is the extent to which people's scores are correlated with other variables or criteria that reflect the same construct.
- **Example:** An IQ test should correlate positively with school performance.
- An occupational aptitude test should correlate positively with work performance.

Two subtypes:

- **Concurrent Validity**: Indicates whether the measurement tool aligns with another valid tool applied at the same time. – IQ Tests
- **Predictive Validity:** Examines the ability of the measurement tool to predict future behaviors or outcomes. – TYT/AYT Exams = Academic Success in the university

Internal Validity (İç Geçerlik)

- Internal validity- Whether the results of a study are truly due to the manipulated independent variables rather than other external factors.
- This is achieved by **controlling potential confounding factors** that could affect the results…
- Internal validity – SO IMPORTANT! - especially in experimental studies, is achieved by eliminating all uncontrolled variables within the study.
- For instance, sleep duration on exam performance…

Dividing students into two groups in a lab. environment:

- **Experimental Group**: Students who sleep 10 hours before the exam.
- **Control Group**: Students who sleep 3 hours before the exam.

Confounding Variables/Factors:

- Study time before the exam,
- Stress levels during the day,
- Dietary habits,
- Previous academic performance…

External Validity (Dış Geçerlik)

- The generalizability of the study
- Findings of a study can be generalized to different populations, settings, and conditions outside the study environment (e.g.lab.)
- Indicates whether the findings from the research apply to **real-life situations.**

- What about previous example?
- «Sleep duration on exam performance»

Other Types of Validity

External Validity Subtypes:

- Ecological validity: How well research findings apply to real-life situations and natural settings.
- Studies conducted in controlled laboratory environments may lack ecological validity - they might overlook various factors that influence behavior in real-world contexts.
    - For example, a study on social interactions conducted in a lab might not yield the same results as it would in participants' everyday environments. High ecological validity ensures that findings remain applicable in real-life settings.

1. High Ecological Validity: The study closely resembles real-life conditions – easy to generalize the results to real-world…
2. Low Ecological Validity: The study environment is too artificial or controlled- difficult to generalize the results to real-world…

Other Types of Validity

Population validity: The findings of a study can be generalized from the sample used in the study to population.

- How well the sample represents the population.
- For instance, if a study includes only college students, the results may not generalize well to individuals of different age groups or educational backgrounds.

Reliability and Validity

Reliable
Not valid

Low validity
Low reliability

Not reliable
Not valid

Both reliable
and valid

Any questions???