Lecture starting soon

Lecture 9: Validity and Reliability

STATISTICS Quantitative Data Analysis in Applied Linguistics



Moustafa Amrate

Department of English, University of Biskra moustafa.amrate@univ-biskra.dz

Outline

• What is validity?

SUM

- Types of validity
- Measuring validity
- What is reliability?
- Types of reliability
- Measuring reliability
- Examples

What is validity?



What is validity?

• **Definition**:

Validity refers to the extent to which a data collection instrument (e.g., **questionnaire**, **test**, **survey**) <u>measures</u> what it is <u>intended to measure</u>.

(Mackey & Gass, 2015, p. 158)



What is reliability?



What is reliability?

• **Definition**:

Reliability refers to the <u>consistency</u> of a measurement instrument, i.e., whether it **yields stable** and **consistent results** across <u>time</u>, <u>items</u>, or <u>observers</u>.

(Mackey & Gass, 2015, p. 180)

Validity vs reliability





Why validity and reliability are important?

- <u>To ensure</u> that the instruments measure what <u>they're supposed to</u> <u>measure consistently</u>.
- To ensure the **t<u>rustworthiness</u>** of results.
- To <u>detect</u> and <u>minimize</u> <u>bias</u> in research.

Types of validity

- 1. Face validity (theoretical)
- 2. Content validity (theoretical)
- 3. Construct validity (theoretical)
- 4. Criterion-Related Validity (practical)

1. Face validity

• **Definition**:

 Face validity refers to the extent to which a test or measurement instrument <u>appears</u> to measure what it claims to measure, at face value.

1. Face validity

• **Definition**:

 Face validity refers to the extent to which a test or measurement instrument <u>appears</u> to measure what it claims to measure, at face value.

• Example:

 You give you pronunciation test to a Phonetics teachers to see if it appropriate/ or if it can measure pronunciation. (Expert judgments)



2. Content validity

• Definition:

Ensures the instrument **fully covers** the **construct being measured** (e.g., <u>all aspects of language proficiency are tested</u>).



2. Content validity

• **Definition:**

Ensures the instrument fully covers the construct being measured (e.g., all aspects of language proficiency are tested).

• Example:

You give your pronunciation test to a **Phonetics teachers** to ensure that **the test covers all elements** of pronunciation (phonemes, stress, intonation ...etc.). (**Expert judgments**)



3. Construct validity

• **Definition**:

Assesses whether the instrument truly measures **the theoretical construct**.



3. Construct validity

• **Definition:**

Assesses whether the instrument truly measures <u>the theoretical</u> <u>construct</u> it intends to measure.

• Example:

If a test claims to measure "vocabulary" an expert would examine whether the test captures the multi-dimensional aspects of vocabulary (e.g., vocabulary inventory, accuracy, retention, appropriateness). (Expert judgments)

4. Criterion-Related Validity

- <u>Definition</u>:
- It refers to the <u>real-world</u> or <u>practical accuracy</u> of the test or data collection instrument.
- Assessing criterion-related validity involves 1) Concurrent Validity and
 2) Predictive Validity.

4. Criterion-Related Validity

1) Concurrent Validity

 Compares the instrument's results to an established test results (benchmark - معيار) collected at the same time.

Example:

High scores in pronunciation scores would **correlate** with success in the **IELTS speaking test** . (Correlation test r/P > 0.75)



4. Criterion-Related Validity

2) Predictive Validity.

Assesses whether the **instrument results** <u>predicts</u> future performance in a **test on related outcomes**.

Example:

High scores in pronunciation test would <u>predict</u> future success in the **US visa oral interview**. (Correlation test *r/P > 0.75*)

Types of validity

- 1. Face validity (theoretical) (Expert judgments)
- 2. Content validity (theoretical) (Expert judgments)
- 3. Construct validity (theoretical) (Expert judgments)
- 4. Criterion-Related Validity (practical) (Correlation test)



What is reliability?

• **Definition**:

Reliability refers to the <u>consistency</u> of a measurement instrument, i.e., whether it **yields stable** and **consistent results** across <u>time</u>, <u>items</u>, or <u>observers</u>.

(Mackey & Gass, 2015, p. 180)





You actually weight 60kg





You actually weight 60kg







You actually weight 60kg



Types of reliability

- 1. Test-Retest Reliability (tests)
- 2. Internal Consistency (tests, survey/ structured questionnaires)
- 3. Inter-Rater Reliability (assessment of test outcomes)

1. Test-Retest Reliability

• **Definition**:

Consistency of results over time when the same instrument is administered **twice under similar conditions**.

Example:

You conduct the same test with two groups of participants of the same proficiency level and under the same conditions. (Correlation test r/P)

2. Internal Consistency

• **Definition**:

Measures the extent to which items within the same instrument are consistent.



2. Internal Consistency (reliability)

- **Definition:**
- Measures the extent to which items within the same instrument are consistent.

Example: (Cronbach's Alpha $\alpha = > 0.75$ Acceptable reliability)

	LIBRARY USE	STRONGLY DISAGREE	NEUTRAL	STRONGLY AGREE
1.	I frequently visit the library to access	0	0	0
	language learning resources			2000
2.	I use library materials to improve my second	0	0	0
	language skills.	Consol		
3.	The library plays an important role in	0	0	0
	supporting my second language learning.		A Second Second	

3. Inter-Rater Reliability

• **Definition**:

Assesses the **agreement** between different observers or **raters** evaluating the **same construct**.

• Example:

Two Phonetics teachers **provide similar** pronunciation scores to the same learner. (Cohen's Kappa *k* = > 0.75 / Percentage agreement)

Validity vs reliability



QUESTIONS?



To be a good researcher is...

- To be curious ... love to learn
- To be critical and logical ... question everything
- To be persistent ... challenges are part of research/life
- To be ethical ... do the right thing even when no-one is looking

THANK YOU!



 \mathbb{X}

 $R^{\scriptscriptstyle \mathsf{G}}$

moustafa.amrate@univ-biskra.dz

@MoustafaAmrate

Moustafa-Amrate