

Chapter 5.
Reliability of
Assessment Device

Contents

- Concept of reliability
- Stability reliability
- Alternate-form reliability

Concept Of Reliability

❖ **Assessment Reliability**

- ✓ *The **consistency** of results produced by measurement devices.*

❖ **Relationship between reliability and validity**

- ✓ *Test reliability is a **necessary**, but **insufficient**.*
Ex) vocabulary test & Pole-Vaulting aptitude test

Stability Reliability

❖ Stability Reliability

- ✓ The *consistency* of assessment results over time.

Appropriate
interval

Administer a
test to
students

Readministe-
r the same
test to the
same
students

Correlate
students'
scores on the
two tests

Result
reliability
coefficient

Stability Reliability

	test		retest
A	7	→	7
B	5	→	6
C	9	→	9
D	8	→	7
E	6	→	6



Stability Reliability

❖ Attention !

✓ *The length of the interval between two test*

The selection of the between-testing interval's length should be made so as to reduce the influence of the first testing on the second, but also to reduce the likelihood that intervening events in the lives of the students will distort the second set of test results.

*'two or three weeks' is reasonable **

Stability Reliability

❖ Attention !

✓ *Same motive, attitude, atmosphere*

Care must be taken to create an atmosphere in which students will devote comparable zeal to both test administrations.

✓ *The length of the test*

the longer the test, the more reliable the test tends to be.

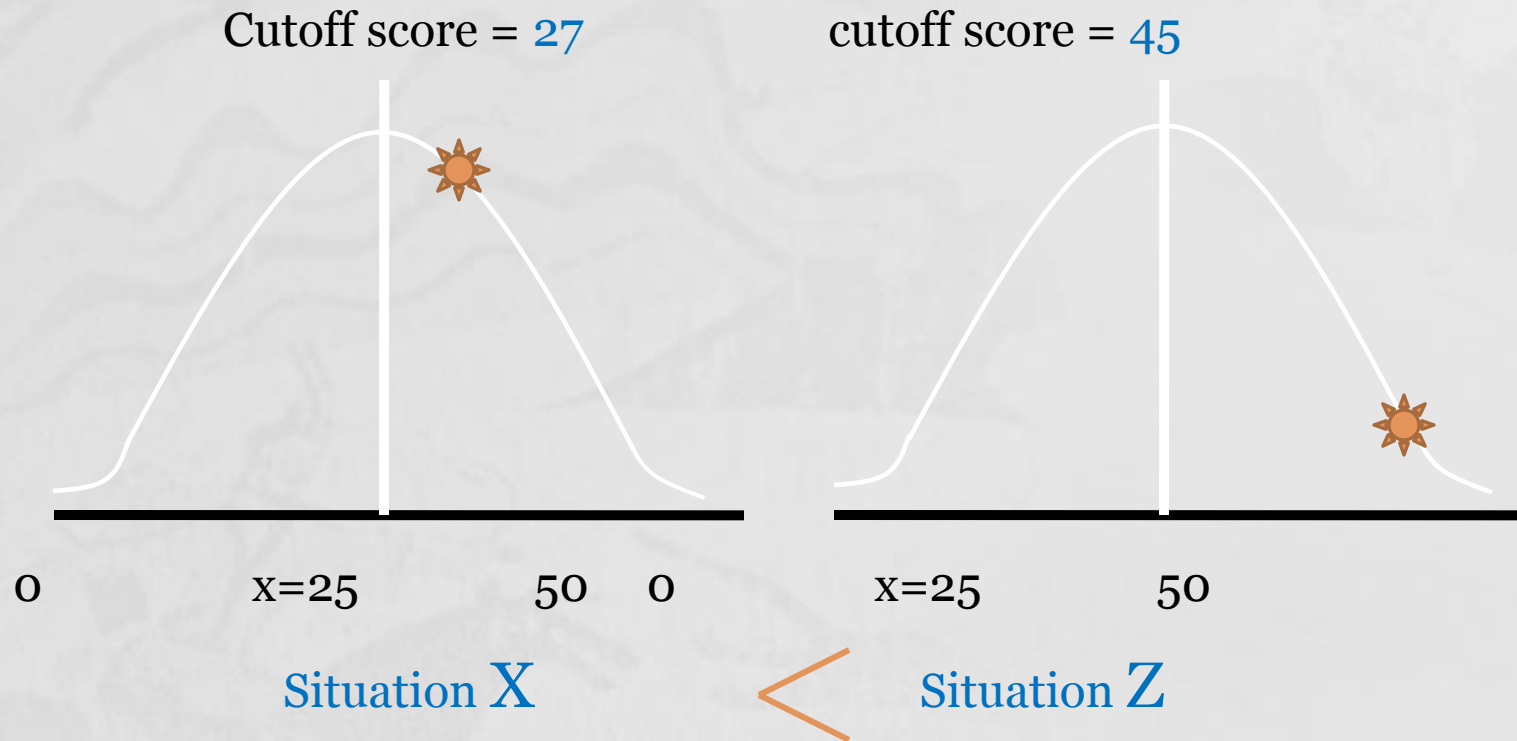
Stability Reliability

❖ Classification Consistency

- ✓ *If the test were administered at a different time, how many students would be classified in the same manner.*
- ✓ *A representation of the **proportion** of students who are placed in the **same category** on two testing occasions or two test forms.*

Stability Reliability

❖ Classification Consistency



Stability Reliability

❖ Classification Consistency

✓ False positive

Classifying a student as *having mastered* a given skill or knowledge domain, but the student actually *had not* mastered it.

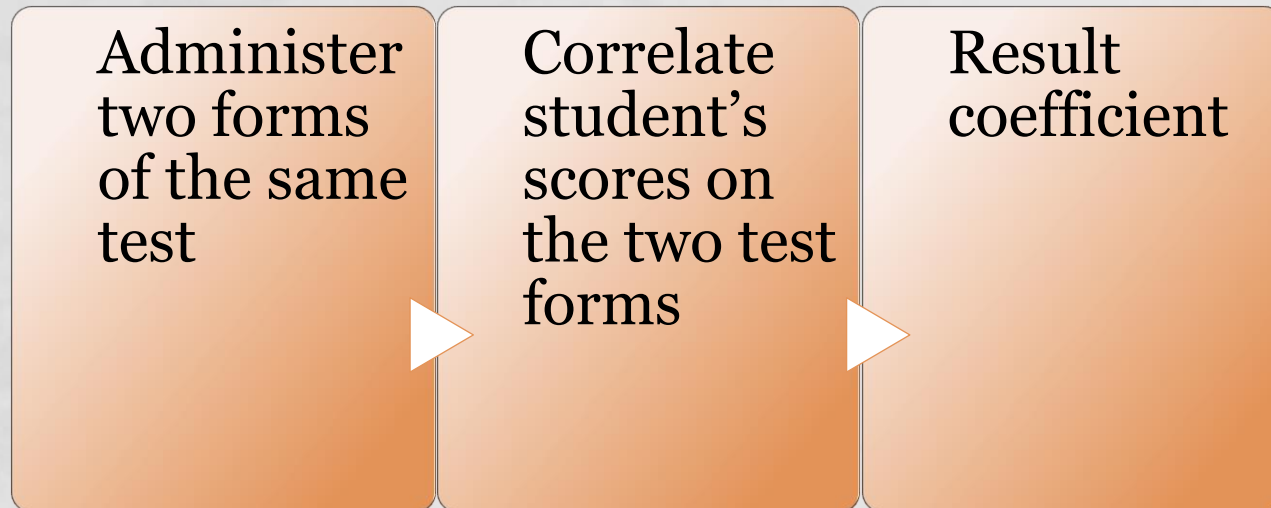
✓ False Negative

Classifying a student as *not having mastered* a given skill or knowledge domain, but the student actually *had* mastered it.

Alternate-Form Reliability

❖ Alternate-Form Reliability

- ✓ *The consistency of measured results yielded by different forms of the same test.*



Alternate-Form Reliability

- ✓ *Reduce remembrance effect (+)*
- ✓ *Difficulty of making different forms of the same test (-)*

Alternate-Form Reliability

❖ Stability and Alternate-Form Reliability

- ✓ *The consistency of measured results **over time** using **two different test forms**.*
- ✓ *has **low** reliability coefficient*
 - ☞ *Because of two factors: time delay and differences between the two forms*

Alternate-Form Reliability

❖ Post Facto Form Equating

- ✓ *two approaches to determining item difficulties*

Field-testing

- : create large numbers of items → administer those items in a formal field test → consequence of field tests → determine the difficulty level of each item*

Embedding

- : embedding of new test items in administered versions of an ongoing testing program*



Thank you.