Dexter M. Balajadia, Ph.D.
Professor, School of Education
University of the Assumption
PHILIPPINES

Session 12

C-EDUC5 (3 units)

ASSESSMENT in Learning 1

Fall Semester 2023

Session 12

C-EDUC5
ASSESSMENT
in Learning 1

Fall Semester
2023

Reliability Testing through Item Analysis of Test Results

Dexter M. Balajadia, Ph.D.

Professor, School of Education
University of the Assumption
P H I L I P P I N E S

Looking back...

Elements of the Table of Specifications (TOS)

Lesson objectives or competencies	Basis of the test
Duration & coverage of instruction	Item proportion
Skills to be assessed	Test items
The type of test	How learners will respond
Instrument design	Overall mechanics

Looking back...

Principles in Planning for the Test

1. A good test is based on the objectives or content areas of the lessons.
2. The number of test items per topic corresponds to the time consumed in class.
3. An efficient test should be challenging enough for students to think critically.
4. Items across topics covered should be randomly placed in the test paper.
5. All test types in the assessment tool are organized in one plan.

Looking back...

Types of TOS

ONE-WAY

Planning a formative test

Table of Specifications of the Formative Test in Math 6

Learning Content	No. of Hours	%	No. of Items	Type of Test and Placement
Place Values of Whole Numbers	3	50	10	Identification 1-10
Rounding Off Numbers	3	50	10	Computation 11-20
Total	**6**	**100**	**20**	

Looking back...

Types of TOS

TWO-WAY Planning a summative test

Table of Specifications of the Summative Test in Math 6

Learning Content	No. of Hours	%	No. of Items	Item Placement Across Skills					
				Remembering	Understanding	Applying	Analyzing	Evaluating	Creating
Numeration Systems	3	12.5	6	MC: 1	MC: 5,6	MC: 16	MC: 18	MC 24	
Whole Number Concepts	6	25	13	MC: 2		MC: 8	MC: 7		PP: 41-50
Ordinal Numbers	2	8	3	MC: 4	MC: 9	MC: 17			
Place Values with Rounding Off	6	25	15		MC: 3	MC: 10,11	MC: 12, 21	PS: 31-40	
Units of Measurement	4	17	8	MC: 13		MC:19	MC: 22 CN: 26-30		
Estimation	3	12.5	5	MC: 14	MC: 15		MC: 20, 23	MC: 25	
Total	**24**	**100**	**50**	**5**	**5**	**6**	**12**	**12**	**10**

MC – Multiple-Choice CN – Conversion PS – Problem-Solving PP – Problem-Posing

Session 11

The Table of Specifications

CONCLUSION

The Table of Specifications (TOS) is the blueprint of the assessment instrument to be crafted. It ensures the validity and quality of the planned test, which indicates the learning content and skills to be assessed and how they are assessed.

Springboard question to ponder

How can we be sure that the test results represent the real quality of learning?

Image 1: Investing in education is a priority for most Filipinos (Source: Basillote, Rao, Martinez, Amoranto & Roque, 2018: Online)

Session 12

Reliability Testing through Item Analysis of Test Results

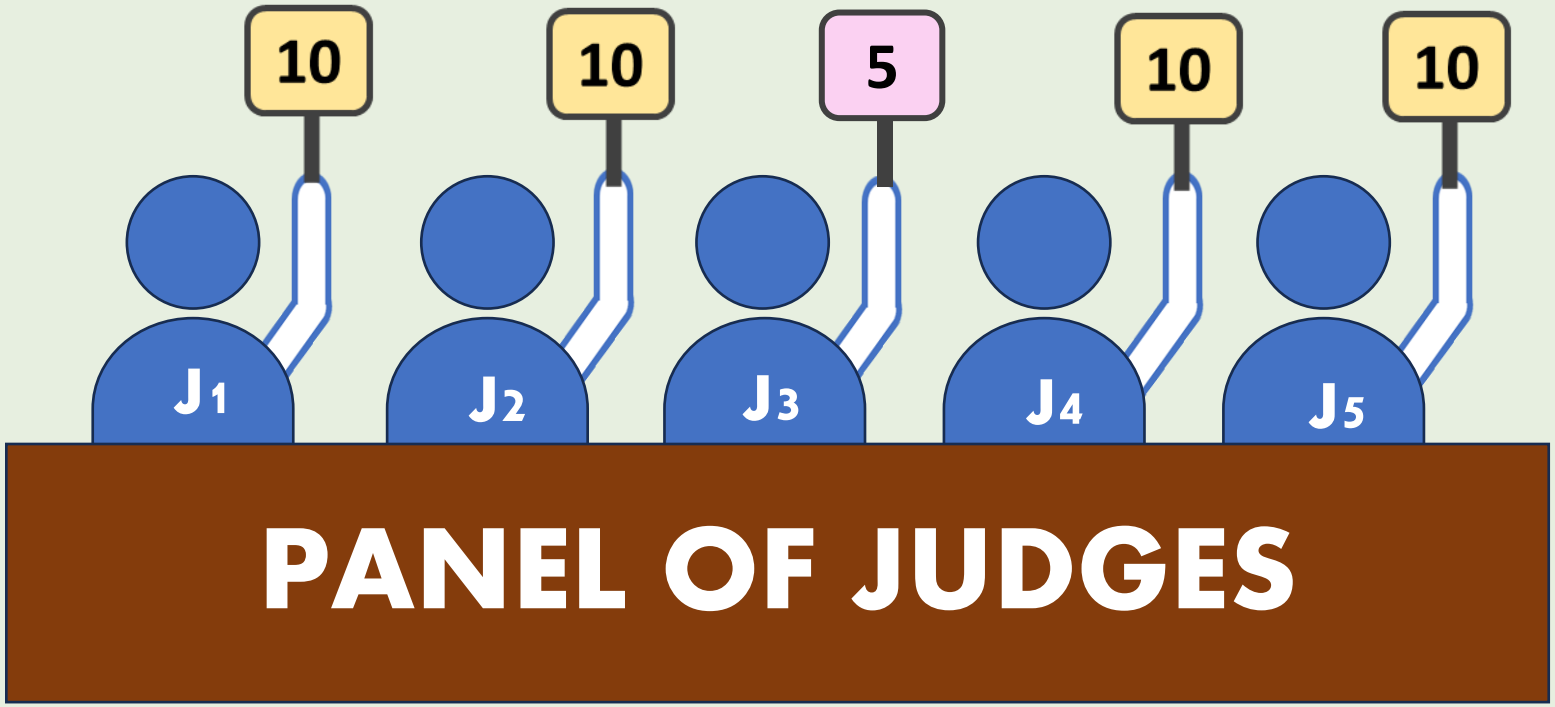
OBJECTIVES

1. Define item analysis and its function in the assessment of learning; and,
2. Determine the three indices in analyzing test item results to describe the quality of teaching and learning.

Topic 1

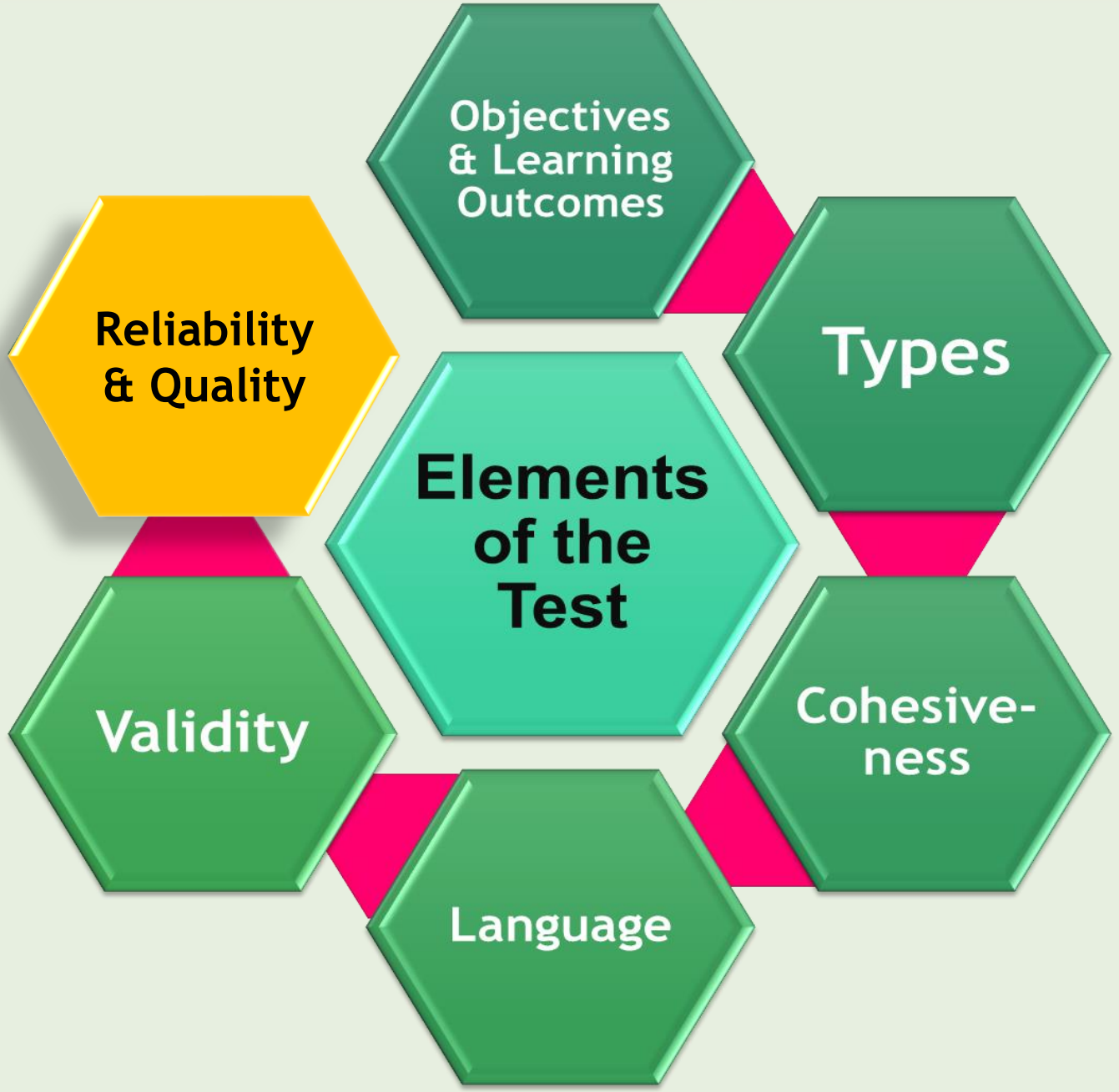
Checking the dependability of test results

What can you say about the scores?



Topic 1

Checking the dependability of test results



Topic 1

Checking the dependability of test results

Reliability & Quality

Are the tests we give to our students able to obtain RELIABLE results?

Ability of the test to elicit consistent responses from the students that reflect their actual performance level

For Multiple-Choice Tests

Analyzing the Quality of the Test Items

Topic 1

Checking the dependability of test results

IMPORTANCE

Providing dependable inputs in planning, implementing, and evaluating the curriculum

Topic 1

Checking the dependability of test results

For Multiple-Choice Tests

Analyzing the Quality of the Test Items

- Item Difficulty
- Discriminatory Power
- Distracting Ability

Topic 1

Checking the dependability of test results

Item Difficulty

The ability of the test item to challenge the student's analytical and critical thinking in the test

Topic 1

Checking the dependability of test results

Discriminatory Power

The ability of the test item to distinguish who really learned and who did not

Topic 1

Checking the dependability of test results

Distracting Ability

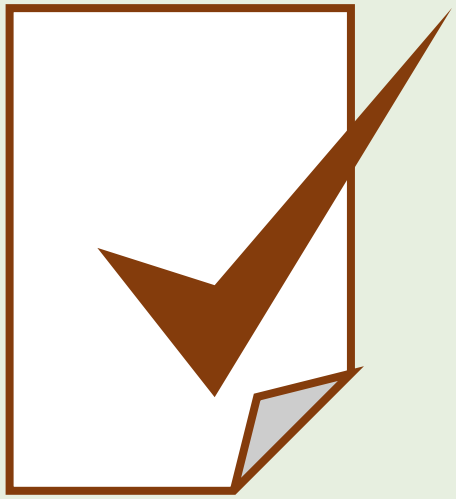
The ability of the wrong or less important options to lure the students who do not know and challenge those who actually know the answer

Item Analysis

Topic 2

Analyzing the quality of the test items

How do we verify if the test items are of good quality?



Topic 2

Analyzing the quality of the test items

Item Analysis

Procedure based on Navarro, Santos and Corpuz (2019):

1. Arrange the test papers from highest to lowest scorers after checking and scoring them.
2. Classify the test papers by taking the 25% of the class who got the highest scores and another 25% who got the lowest scores.

Item Analysis

Procedure based on Navarro, Santos and Corpuz (2019):

3. Make a worksheet to record and prepare the data for analysis.

Topic 2

Analyzing the quality of the test items

Item No.	Highest Scorers (nH=12)				Lowest Scorers (nL=12)			
	A	B	C	D	A	B	C	D
1	1	8	1	2	4	4	2	2
2	0	5	2	5	3	6	1	2
3	3	3	3	3	5	3	2	0
4	0	0	0	12	2	0	2	8
5	4	2	6	0	3	5	3	1

Topic 2

Analyzing the quality of the test items

Item Analysis

Procedure based on Navarro, Santos and Corpuz (2019):

4. Prepare to compute the P for each item by doing the following:

$$N = nH + nL$$

Where:

N – Total Number of Students considered
nH – number of highest-scoring students
nL – number of lowest-scoring students

$$N = 12 + 12 = \underline{24}$$

Item Analysis

Procedure based on Navarro, Santos and Corpuz (2019):

Topic 2

Analyzing the quality of the test items

5. Compute the difficulty index (P) of every test item using this formula:

$$P_i = \frac{cH + cL}{N}$$

P_i – Difficulty index per item

cH – Number of students in the highest group who got the correct answer

cL - Number of students in the lowest group who got the correct answer

N – Total number of students considered

Item Analysis

Procedure based on Navarro, Santos and Corpuz (2019):

Item No.	Highest Scorers (nH=12)				Lowest Scorers (nL=12)			
	A	B	C	D	A	B	C	D
1	1	8	1	2	4	4	2	2

$$P_i = \frac{cH + cL}{N} \Rightarrow P_i = \frac{8+4}{24} = \frac{12}{24} = 0.50$$

Range	Interpretation	Action
0.00 – 0.25	Difficult	Discard
0.26 – 0.75	Moderately Difficult (Ideal result)	Retain
0.76 and above	Easy	Revise

Topic 2

Analyzing the quality of the test items

Item Analysis

Procedure based on Navarro, Santos and Corpuz (2019):

Compute the Ps of the other items.

Item No.	Highest Scorers (nH=12)				Lowest Scorers (nL=12)			
	A	B	C	D	A	B	C	D
1	1	**8**	1	2	4	**4**	2	2
2	0	5	**2**	5	3	6	**1**	2
3	**3**	3	3	3	**5**	3	2	0
4	0	0	0	**12**	2	0	2	**8**
5	4	**2**	6	0	3	**5**	3	1

P – Interpretation
0.50 – Mod. Difficult
0.13 – Difficult
0.33 – Mod. Difficult
0.83 – Easy
0.29 – Mod. Difficult

Topic 2

Analyzing the quality of the test items

Item Analysis

Procedure based on Navarro, Santos and Corpuz (2019):

6. Compute the discriminatory index (D) of every test item using this formula:

$$D_i = \frac{cH - cL}{n}$$

Where:

D_i – Discriminatory index per item

cH – Number of students in the highest group who got the correct answer

cL – Number of students in the lowest group who got the correct answer

n – 25% of the class

Item Analysis

Procedure based on Navarro, Santos and Corpuz (2019):

Item No.	Highest Scorers (nH=12)				Lowest Scorers (nL=12)			
	A	B	C	D	A	B	C	D
1	1	8	1	2	4	4	2	2

$$D_i = \frac{cH - cL}{n} \Rightarrow D_i = \frac{8-4}{12} = \frac{4}{12} = 0.33$$

D value range	Interpretation
-1.00 to -0.60	Questionable
-0.59 to -0.20	Not Discriminating
-0.19 to 0.20	Moderately Discriminating
0.21 to 0.60	Discriminating (Ideal Result)
0.61 to 1.00	Very Discriminating (Most Ideal Result)

Topic 2

Analyzing the quality of the test items

Item Analysis

Procedure based on Navarro, Santos and Corpuz (2019):

Compute the Ds of the other items.

Item No.	Highest Scorers (nH=12)				Lowest Scorers (nL=12)			
	A	B	C	D	A	B	C	D
1	1	**8**	1	2	4	**4**	2	2
2	0	5	**2**	5	3	6	**1**	2
3	**3**	3	3	3	**5**	3	2	0
4	0	0	0	**12**	2	0	2	**8**
5	4	**2**	6	0	3	**5**	3	1

D – Interpretation

0.33 – Discriminating

0.08 – Mod. Disc'g.

-0.17 – Mod. Disc'g.

0.33 – Discriminating

-0.25 – Not Disc'g.

Topic 2

Analyzing the quality of the test items

Deciding about the item

Item Analysis Procedure based on Navarro, Santos and Corpuz (2019):

Difficulty Level	Discriminating Level	Decision
Difficult / Very Difficult	Not Discriminating / Questionable	Improbable – Discard
	Moderately Discriminating	May need revision on the stem and/or choices
	Discriminating / Very Discriminating	Accept with little revision if necessary
Moderately Difficult	Not Discriminating / Questionable	Needs revision especially on the choices
	Moderately Discriminating	Accept but needs slight revision on the stem or choices
	Discriminating / Very Discriminating	Accept as it is
Easy / Very Easy	Not Discriminating / Questionable	Totally discard
	Moderately Discriminating	Needs major revision on the stem and/or choices
	Discriminating / Very Discriminating	Accepted by slightly increasing difficulty

Item Analysis

Procedure based on Navarro, Santos and Corpuz (2019):

7. Decide the fate of the test items based on their P and D.

Item Number	P – Interpretation	D – Interpretation	Decision
1	0.50 – Mod. Difficult	0.33 – Discriminating	1. Accept
2	0.13 – Difficult	0.08 – Mod. Disc'g.	2. Revise
3	0.33 – Mod. Difficult	-0.17 – Mod. Disc'g.	3. Revise/Discard
4	0.83 – Easy	0.33 – Discriminating	4. Improve
5	0.29 – Mod. Difficult	-0.25 – Not Disc'g.	5. Discard

Topic 2

Analyzing the quality of the test items

Item Analysis

Procedure based on Navarro, Santos and Corpuz (2019):

8. Counter-check with the distracters.

An option is ineffective when no one chose it.

Topic 2

Analyzing the quality of the test items

Item Analysis

Procedure based on Navarro, Santos and Corpuz (2019):

8. Counter-check with the distracters.

The option is good when more in the lowest group chose it than those who did in the highest group.

Topic 2

Analyzing the quality of the test items

Item Analysis

Procedure based on Navarro, Santos and Corpuz (2019):

8. Counter-check with the distracters.

An option chosen by less than 10% of the class in an easy item is not an effective distracter.

Topic 2

Analyzing the quality of the test items

Item Analysis

Procedure based on Navarro, Santos and Corpuz (2019):

8. Counter-check with the distracters.

In a questionable or not discriminating item, the most chosen distracter is misleading.

Session 12

Reliability Testing through Item Analysis of Test Results

CONCLUSION

With item analysis, you can see how each student in your class did on each question. The Difficulty Index, the Discriminatory Index, and the Distracter Analysis are three popular types of item analysis that give teachers distinct kinds of information for decision-making in the teaching-learning process.

Session 12

Reliability Testing through Item Analysis of Test Results

REFERENCES

Basillote, L., Rao, L.N., Martinez, A., Amoranto, G. & Roque, J.D. (2018). *Investing in education is a priority for most Filipinos* [Online Image] [Accessed on October 30, 2023] https://blogs.adb.org/blog/5-ways-make-most-philippine-education-investments

Kubiszyn, T. & Borich, G. (2007). *Educational testing and measurement: Classroom application and practice. 8th edition.* N.J., U.S.A.: John Wiley & Sons, Inc.

Navarro, R.L., Santos, R.G., & Corpuz, B.B. (2019). *Assessment in learning 1. 4th edition.* Quezon City, Philippines: LORIMAR Publishing