

Session 11

Engineering the Test: The Table of Specifications (TOS)

A. Introduction

Validity in evaluation relies heavily on alignment as a fundamental component. In order to boost the students' levels of motivation and encourage the development of effective learning practices, we need to make sure that our pupils are aware of this alignment. After all, we cannot accurately evaluate our children's skills if the emphasis placed on their education is wrong. Let's look closer at aligning our assessments with our learning objectives and activities to ensure we assess what we plan to measure.

Our previous lectures emphasized the importance of test validity and how to create valid test items. The standards serve as a helpful reminder that the intended use of the results is equally as important as the test's characteristics when considering validity.



Image 1: Grade school students in the Philippines are lagging behind counterparts in Southeast Asia in reading, math and writing, a study showed (Source: De Guzman, 2020: Online)

Content validity shows if students met course learning goals. When test questions correspond to course goals and learning activities at the correct cognitive level (Bloom's Taxonomy), they are valid. We check construct validity to evaluate if our assessment is academically sound. According to our theory, construct validity is supported by test results, which measure something unseen, being tightly linked to other variables. Criteria validity predicts our pupils' external exam results. Significantly similar exam scores show criteria validity. Face validity ensures the test's physical, visual, and performance integrity.

Session 10 Conclusion

The test item is valid when the content is based on the learning outcome being assessed and is proven to measure the learning skill intended in its construction. Besides the content and construct validity, the entire test's physical features and language proficiency make face validity equally essential.

B. Session Objectives

This lecture will facilitate your accomplishment of the following:

1. To consider the teaching-learning elements in planning for the test
2. To configure the teaching-learning elements into a table of specifications of a summative test

C. Session Content

1. Teaching and Learning for Assessment

The teacher needs to check learning and determine whether students learn what and how they should learn. When it comes to developing the structure of the class evaluation, this is one of the most essential questions that need to be considered. Imagine an assessment plan for a class as a matrix that maps formative and summative evaluations against your learning goals and integrates qualitative and quantitative data in variable amounts.

Formative evaluations are opportunities to practice learning tied to a lesser consequence than summative exams. Formative assessments might take the shape of homework assignments, in-class quizzes, collaborative projects, or drafts of papers. Students can use formative assessments, which are lower-stakes evaluation chances, to scaffold their learning as they progress toward summative assessments, opportunities for high-stakes evaluation such as examinations, and extensive or conclusive projects. Students can better prepare for summative examinations by first completing formative assessments.

The formative and summative assessments administered to students should be combined to allow them to study, practice, and demonstrate their progress toward the course's learning goals. In this regard, teachers or test constructors must consider a systematic way of planning assessment, especially summative and achievement tests.

Five main elements of the assessment plan are considered to ensure the validity and quality of the instrument to be used in assessing the teaching and learning process. These are (a) lesson objectives or competencies, (b) duration or coverage of instruction, (c) skills to be assessed, (d) the type of test, and (e) instrument design.

a) Lesson objectives or competencies

Based on the constructive alignment in the educational curriculum, lesson objectives, learning outcomes, or competencies are the bases of instruction and assessment. Test items are derived from these goals as how the teaching-learning experiences have been designed from them.

b) Duration and coverage of instruction

Fairness and balance are two qualities a credible assessment tool must possess. The number of items or the credit score in a test is based on the duration of the teaching-learning experiences. The skills where students dwelled the most have the most share of the test proportion.

c) Skills to be assessed

Frequently, lesson objectives require more than one skill to be tested, no matter how they are expressly stated. In this case, more than one test item is constructed to measure learning reliably. Therefore, the test constructor also designs the test according to the various skills intended for the particular learning outcome.

d) The type of test

In choosing the type of test to be constructed, the teacher looks at the lesson objective side-by-side with the skills to be assessed. These fundamental elements determine how students are evaluated and how they should demonstrate their learning. The type of test determines how students should respond and what to respond.

e) Instrument design

Aside from requirements for face validity of the test, the instrument design is another essential consideration. The design covers the types and layout and how the items are arranged. Typically, they are arranged from the easiest to the most difficult; however, items under the same objective are scattered randomly in the instrument.

2. The Table of Specifications (TOS)

The table of specifications (TOS) aims to guarantee that the exam or assessment measures the desired knowledge and abilities. Therefore, it can give evidence for both the content and construct validity of responses when applied appropriately. Using a TOS to create large-scale tests, conduct teacher evaluations, and create psychometric scales is possible. It's a must-have for any researcher or teacher interested in creating their own tests or measurements.

A TOS is typically created to guarantee that assessment items and other parts measure the correct knowledge, abilities, and attitudes. To put it another way, a test content issue focus provided by a TOS aids test constructors in ensuring that the test or assessment measures the desired constructs. It would be reasonable for a teacher to include a task in which pupils draw the moon's phases to gauge their comprehension of lunar phases. But if you wanted to evaluate students' understanding of lunar phases, you couldn't use an item that asked them to name the first person to set foot on the moon.

A TOS can also demonstrate response process validity for test constructors. The response method is how test takers should think when taking the test. A teacher may expect students to memorize the lunar phases; therefore, a knowledge-level inquiry is acceptable. If the teacher teaches lessons that involve tracking the moon, creating journals, and discussing its phases, the assessment should focus on higher-level thinking skills like analysis, evaluation, and synthesis. Incorporating annotated explanations into a lunar phase model might better align with students' ideas during education.

This lecture will guide you in preparing the TOS required in all Philippine schools, especially in crafting summative, quarterly, periodical, major, and other achievement tests.

There are two acceptable types of TOS according to the purpose and scope of the planned assessment.

One-Way TOS. This type of TOS is usually done in planning formative and short assessments. Only four of the five elements of the assessment plan are exhibited. The learning skills to be tested are not specified in the one-way TOS, as shown in Table 1.

Two-Way TOS. This type is more recommended or required in most schools because it is complete with all the elements of the assessment plan. The two-way TOS shows how test items are placed in the instrument and the learning skills they measure; hence, it is two-way (see Table 2).

Table 1. Sample One-Way TOS

Table of Specifications of the Summative Test in Math 6

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

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

Table 2. Sample Two-Way TOS

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

Steps in Constructing the TOS (Two-Way)

1. Prepare a table with ten columns whose column widths will be adjusted according to the length of data to be filled in.
2. Specify the column headings or titles, to wit:

| Learning Content | No. of Hours | % | No. of Items | Item Placement Across Skills | | | | | |
|------------------|--------------|---|--------------|------------------------------|---------------|----------|-----------|------------|----------|
| | | | | Remembering | Understanding | Applying | Analyzing | Evaluating | Creating |

3. Enumerate the “Learning Contents” in the first column. Some schools require lesson objectives, learning outcomes, competencies, or topics instead of learning content.
4. Indicate the duration of each learning content according to the actual interaction in the second column. You may express the data as “No. of Hours, Minutes, Days, Weeks, etc.” depending on the school standards. Customarily, the duration of instruction is based on the syllabus or course curriculum. Indicate the total duration at the bottom row.
5. Compute the proportion (%) of the duration of instruction (DI) per learning content by the total duration (TD). You may round off the results to the whole number if necessary. Make sure that the total is always 100%.

$$\% = \frac{DI}{TD} \times 100$$

6. In the fourth column, determine the total number of items or points of the test (TI). Compute the number of test items or points (NI) for every learning content by getting the percentage of the total items.

$$NI = TI \times \%$$

7. Adjustment in the number of items per learning content is allowed so that the teacher can flexibly redistribute the items based on the possibility of creating test items. A difference of two to five between the computed percentage and the adjusted quantity is expected. However, the test constructor should exercise fairness according to the actual duration or emphasis of the learning skill.
8. Based on the learning content, learning outcome, and the number of items, determine the types of tests to be used. Code them right below the table as legends (e.g., MC – Multiple-Choice, CN – Conversion, PS – Problem-Solving, PP – Problem-Posing, etc.).

9. Indicate the six cognitive skills across the remaining six columns according to Bloom's Taxonomy – Remembering, Understanding, Applying, Analyzing, Evaluating, and Creating. Place the category title as "Item Placement Across Skills".
10. Distribute the item numbers chronologically from item 1 per learning content and cognitive skill. When doing this, keep an eye on the total number of items per topic. Put the test type code before the corresponding item numbers.
11. Get the total number of items per cognitive skill. There should be more higher-order test items (analyzing, evaluating, creating) than lower-order ones (remembering, understanding, applying) to make a high-quality summative test. In this lecture, it is recommended to adopt between 60%:40% and 80%:20% higher to lower-order, respectively.
12. Double-check and proofread the TOS before making the test items.

D. 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.

E. References

- De Guzman, M. (2020). *Grade school students in the Philippines are lagging behind counterparts in Southeast Asia in reading, math, and writing, a study showed* [Online Image] [Accessed on October 27, 2023] <https://www.bworldonline.com/editors-picks/2020/12/03/331914/filipino-students-falling-behind-in-reading-writing-levels-in-southeast-asia/>
- 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