

## Measurement of Learning Outcomes

-----

### 1. Definition of Measurement and Evaluation

- **Measurement:** The process of assigning numbers or symbols to characteristics of what is being measured according to specified rules. Examples include giving scores on a test (e.g., 1 or 0), counting the frequency of a behavior, or scoring the completeness of an answer (e.g., 3, 2, 1 based on clarity of response). Measurement is always the necessary first step before evaluation.
- **Evaluation:** The process of using measurement data to make a judgment about value or worth based on predetermined criteria. For example, stating that a score of 28 out of 40 is a "pass" because the criterion is 70%. Formula:  $E = M + J$  (Evaluation = Measurement + Judgment based on specified criteria). Evaluation results in classifications such as excellent/moderate/poor, good/bad, outstanding/needs improvement.

### 2. Measuring Learning Outcomes: 3 Domains

Effective measurement must cover the expected learning outcomes across three main domains:

- **Cognitive Domain:** Measures knowledge, comprehension, analysis, evaluation, and creation.
- **Affective Domain:** Measures attitudes, values, interests, and desirable characteristics.
- **Psychomotor Domain:** Measures practical skills, work processes, and the quality of products or performances.

### 3. Components of Measurement

Measurement requires a target/objective to be measured + a set of indicators (e.g., test questions, items in a psychomotor or affective domain scale) + a scoring scale (e.g., 1-0, 3-2-1, or 5-4-3-2-1).

### 4. Tools Used in Learning Assessment

The chosen tools must align with each learning outcome domain:

- **Cognitive Domain:** Typically uses tests (both objective - e.g., multiple choice, and subjective - e.g., essays, fill-in-the-blank).
- **Affective Domain:** Typically uses attitude scales or classroom behavior observation forms.

- **Psychomotor Domain:** Typically uses work process evaluation forms or product/performance assessment forms. In practice, these are often called "evaluation forms" even though they are essentially measurement tools attempting to quantify skills into numerical levels. They are called evaluation forms because they rely on the estimator's (observer's) judgment.

## 5. Quality Verification of Measurement Tools

### 5.1 Overall Quality Verification of the Tool

After constructing a tool, two key quality dimensions must always be considered together:

- **Validity:** The ability of a tool to accurately measure what it intends to measure. This is the most important property but the most difficult to prove.
- **Reliability:** The consistency and stability of scores across different situations or repeated measurements (2-3 times). Reliability gives the measurer confidence in their conclusions.

### 5.2 Item Analysis

Beyond overall quality, measurement specialists should analyze the quality of each individual item/question:

- **Difficulty Index (p):** To determine if an item is too difficult or too easy.
- **Discrimination Index (r):** To determine an item's ability to distinguish between high-achieving and low-achieving individuals.
- **Item-Total Correlation:** To examine the relationship between the score on a single item and the total test score.