The Initial Steps in Developing a Classroom Test
Careful assessment design is important, as it is estimated that 20-35% of available instruction time is consumed by assessment-related activities.

Learning objectives = Educational goals, also referred to as instructional or learning objectives; a good classroom test can be written much more easily from clearly written objectives.
Characteristics of Objectives

- **Scope**: Refers to whether an objective is specific or broad
- **Domain**: Refers to the cognitive, affective, or psychomotor domain
- **Format**: Refers to behavioral versus nonbehavioral
Scope

- Broad objectives are typically broken down into more narrow ones to support students and guide them to the broader objective, for instance:
  - 1. The student will be able to discuss the effects of the Civil War on twentieth-century American politics.
    - 1a. The student will be able to discuss the political effects of post-Civil War occupation by federal troops on the Southern state politics.
    - 1b. The student will be able to trace the rise and fall of African Americans’ political power during and after Reconstruction.

- Objectives should either be of intermediate specificity, or a good balance of specific and broad.
The three domains typically addressed by educational objectives are cognitive, affective, and psychomotor abilities.

Each of these domains are usually presented as hierarchies with different levels that reflect varying degrees of complexity.
Cognitive Domain

- Cognitive objectives: Deal with memorizing, interpreting, and analyzing
- Bloom’s taxonomy categorizes the complexity of various cognitive objectives
  - Knowledge – remembering
  - Comprehension – understanding
  - Application
  - Analysis
  - Synthesis
  - Evaluation

Cognitive processes underlying these categories is assumed to be increasingly complex, though these are not hierarchical.
## Bloom’s Taxonomy

<table>
<thead>
<tr>
<th>Level</th>
<th>Description</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge</td>
<td>Rote memory, learning facts</td>
<td>Name each figure of speech</td>
</tr>
<tr>
<td></td>
<td>Summarize, interpret, or explain material</td>
<td>Explain each figure of speech.</td>
</tr>
<tr>
<td>Comprehension</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Application</td>
<td>Use general rules and principles to solve new problems</td>
<td>Given a poem, identify different figures of speech.</td>
</tr>
<tr>
<td>Analysis</td>
<td>Reduction of concepts into parts and show the</td>
<td>Compare and contrast the function of different figures of speech</td>
</tr>
<tr>
<td></td>
<td>relationship of parts to the whole</td>
<td></td>
</tr>
<tr>
<td>Synthesis</td>
<td>Creation of new ideas or results from existing</td>
<td>Write a poem, using different figures of speech</td>
</tr>
<tr>
<td></td>
<td>concepts</td>
<td></td>
</tr>
<tr>
<td>Evaluation</td>
<td>Judgment of value or worth</td>
<td>Edit written work that includes the use of figures of speech, using</td>
</tr>
<tr>
<td></td>
<td></td>
<td>evalutive criteria</td>
</tr>
</tbody>
</table>
Using Bloom’s Taxonomy

• Objectives written to tackle higher levels of Bloom’s taxonomy are desirable, but not always realistic.
• For example, when introducing a topic, lower levels may be incorporated into the core objectives of a lesson or unit.
• Complexity of objectives should be built up over time, usually as a result of feedback or reflection.
Affective Domain

- Affective objectives: Involve the attitudes and actions of students in relation to a school subject, for example:
  - The student will demonstrate interest in earth science by conducting a science fair project in some area of earth science.
- Generally emphasized more in elementary school than secondary
- Krathwohl’s taxonomy categorizes the complexity of various affective objectives
Psychomotor Domain

- Psychomotor objectives: Deal with physical activity
- Usually occur in physical education, dance, speech, theater, laboratory, or career-technical classes
- Harrow’s taxonomy categorizes the complexity of various psychomotor objectives
Behavioral versus Nonbehavioral Educational Objectives

- **Behavioral objectives** = Activities that are observable and measurable
  - Typically identify an outcome criterion
  - Examples: Arrange, build, create, define

- **Nonbehavioral objectives** = Activities that are unobservable and not directly measurable (must be inferred)
  - Examples: Analyze, examine, judge, understand
Tips for Writing Learning Objectives

1. Write objectives that cover a broad spectrum of abilities
2. When feasible, identify behaviors that are observable and directly measurable
3. State special conditions
4. When appropriate, specify an outcome criterion (this is particularly necessary for objectives that are to be assessed)
Developing a Table of Specifications

- This essentially serves as your test blueprint
- The first column lists the content areas to be assessed
- The rest of the columns are devoted to levels of the objectives assessed by individual items according to Bloom’s categories
- The body of the table consists of the number of each item in every category
- A final column lists the totals of items in each content area
- This table can then be converted into percentages, which some experts prefer
**Example of a Table of Specifications**

<table>
<thead>
<tr>
<th>Content Area</th>
<th>Knowledge</th>
<th>Comprehension</th>
<th>Application</th>
<th>Analysis</th>
<th>Synthesis</th>
<th>Evaluation</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number line</td>
<td>2</td>
<td>2</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>Addition using Number line</td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>Subtraction, using number line</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
<td>2</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>Addition and subtraction of mixed numbers</td>
<td>2</td>
<td>3</td>
<td></td>
<td>2</td>
<td>2</td>
<td></td>
<td>9</td>
</tr>
</tbody>
</table>
- Ideally, a table of specifications should help you to ensure you cover the curriculum content and have items of varying complexity
- After designing the test, you can determine how you will interpret the scores:
  - Norm-referenced scores= The performance of a student is relative to how other students performed (e.g. percentile ranks and standard scores) – not typical of teacher-made tests
  - Criterion-referenced scores= Absolute, and represent how well the individual student has mastered the content (e.g. percent correct and mastery/nonmastery scores)
Selecting Which Types of Items to Use

- First, consider the objectivity of your test items:
  - Multiple-choice items and true-false and matching items should have a clearly correct answer (it is objectively true as to which options are correct and incorrect)
  - Essays and other test items may be more subjective
Selecting Which Types of Items to Use

Next, consider the type of response you want:

- **Selected-response items** like multiple choice, true-false, and matching items require their answers to be picked from a list of options
- **Constructed-response items** like fill-in-the-blank, short-answer, and essay items are provided entirely by the student, and also include performance assessments and portfolios

Careful consideration of the strengths and weaknesses of each type, along with the nature of the objective you wish to assess, should be given before deciding which types of items to use
Selecting Which Types of Items to Use

- **Strengths of selected-response items**
  - You can have a large number of items on your test, and it may be easier to cover the content domain as a result
  - Can be scored efficiently, objectively, and reliably
  - Does not require as much skill in grading
  - Are good for measuring lower-level objectives
  - Reduce the influence of some construct-irrelevant factors (e.g., writing ability)

- **Weaknesses of selected-response items**
  - Relatively difficult to write
  - Not able to assess all learning objectives
  - Subject to effects of random guessing
Selecting Which Types of Items to Use

- **Strengths of constructed-response items**
  - Often easier and less time-consuming to write compared to selected-response items
  - Well suited for assessing higher-order cognitive abilities and complex task performance
  - Eliminate random guessing

- **Weaknesses of constructed-response items**
  - Take more time for students to complete, so not as many items can be included on a test (difficult to sample content)
  - More difficult to score reliably
  - Vulnerable to feigning
  - Vulnerable to the influence of construct-irrelevant factors
Putting Assessment Together

1. Follow your table of specifications
2. Provide clear directions
3. State the question, problem, or task in as clear and straightforward a manner as possible
4. Develop items and tasks that can be scored in a decisive manner
5. Avoid inadvertent clues to the correct answers
6. Arrange the items in an assessment in a systematic manner (e.g., by objective, type of items)
7. Include test items and tasks that will result in an assessment that produces reliable and valid test results
8. When determining how many items you should have on a test, consider students’ ages, types of items employed, and the type and purpose of the test
Preparing Your Students and Administering the Assessment

- Promote conditions in your classroom under which students can perform their best
- Make sure students know when the test is, what it will cover, the basic parameters (time limit and item types), how it will be scored, and how the results will be used
- Give students example items to practice and instruct them in basic test-taking skills
- Motivate your students to prepare without stressing them out too much; avoid high-stakes outcomes of in-class tests
- Schedule the test carefully (e.g. not immediately before or after a holiday)
- Make efforts to avoid unexpected interruptions
- Clarify students’ questions about the test without dropping hints
- Discourage all forms of cheating