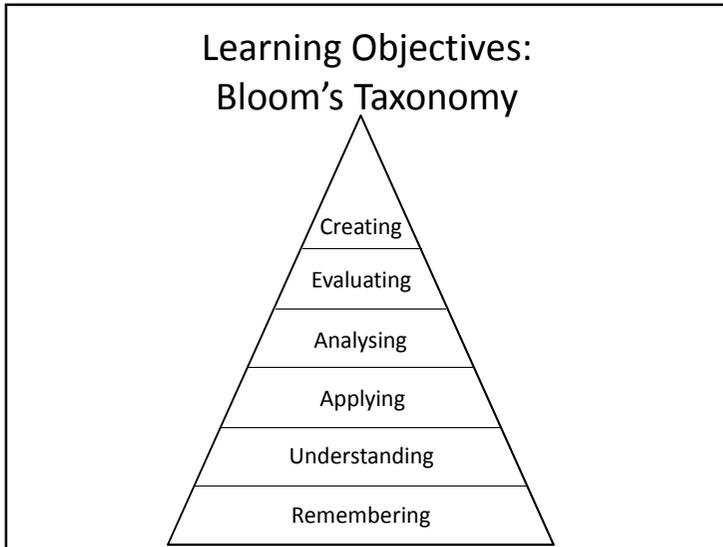


Developing Effective Assessments

Dr. K. A. Korb
 University of Jos

Overview of Classroom Assessment

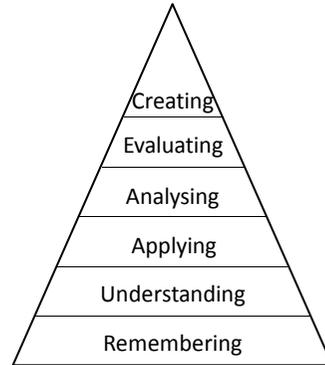
- Based on the curriculum, set objectives for student learning through instruction
- After instruction, assessment is conducted to determine whether objectives have been met
 - If assessment verifies that learning objectives have been met, then continue to the next learning objectives
 - If assessment provides evidence that learning objectives have not met, then additional instruction in those learning objectives are necessary



Category	Definition	Keywords for Objectives
Remember	Recall information	Recite. Define. Identify. Describe. Recognize. Know. List. Name.
Understand	Understand the meaning of a concept	Comprehend. Explain. Summarize. Translate.
Apply	Use a concept in a familiar situation; Use a procedure	Apply. Compute. Predict. Demonstrate. Implement. Use.
Analyze	Break information into parts to explore relationships	Analyze. Compare. Contrast. Differentiate. Distinguish. Diagram.
Evaluate	Make judgment about the value of a concept based on standards	Criticize. Critique. Evaluate. Justify. Support.
Create	Generate new ideas, products, or ways of viewing something	Create. Design. Generate. Plan. Construct. Produce. Invent.

Bloom's Taxonomy

1. What is the definition of rote memorization?
2. What are the advantages and disadvantages of teaching for meaningful learning?
3. How can you use meaningful learning in your classroom?
4. Create your own lesson that focuses on meaningful learning.
5. What are the similarities between meaningful learning and constructivism?
6. Why should students engage in meaningful learning?



Classroom Assessment

- Grading System
 - Clear and understandable
 - Designed to support learning and provide frequent feedback
 - Based on hard data
 - Fair to all students
 - Able to defend to parent and/or administrator

Types of Assessment

- **Objective items:** Specific predetermined answer
 - Multiple Choice
 - True/False
 - Matching
 - Fill in the blank
- **Subjective items:** Teacher makes judgment to assess
 - Essay
 - Short Answer
 - Performance Assessment

Creating Forced-Choice Items

- Present one clear problem in the stem
- Make all distracters probable (but clearly incorrect)
- Vary the position of the correct choice randomly (try to avoid C)
- Avoid similar wording in the stem and correct choice
- Avoid phrasing the correct choice in more technical terms than distracters
- Correct answer and distracters should have same length

Creating Forced-Choice Items

- Avoid absolute terms (always, never)
- Keep stem and distracters grammatically consistent
- Avoid using two distracters with the same meaning
- Emphasize NEGATIVE wording
- Use “none of the above” with care
- Avoid “all of the above”

Bad Forced-Choice Items

- The most significant cause of World War II was:
 - a) American aid to Great Britain
 - b) Italy’s conquering of Ethiopia
 - c) Japan’s war on China
 - d) the devastation of the German economy as a result of the Treaty of Versailles
- Bats are the most common mammal. The one of the following that is not a mammal is an:
 - a) Wolf
 - b) Giraffe
 - c) Alligator
 - d) Leopard

Bad Forced-Choice Items

- Which is a function of the circulatory system?
 - a) to support the vital organs of the body
 - b) to circulate blood through the body
 - c) to transfer nerve impulses from the brain to the muscles
 - d) to provide for the movement of the body’s large muscles
- The definition of population density is:
 - a) how many people live in your city
 - b) how many people voted
 - c) frequency of people per square mile in a country
 - d) how many people are married

Essay Items

- Strengths
 - Assess creative and critical thinking
 - Students more likely to meaningfully organize information when studying
- Weaknesses
 - Scoring takes time
 - Scoring can be biased

Essay Items

- Questions should clearly indicate the elements to be covered in the answer.
 - Bad: Explain operant conditioning.
 - Good: Explain the four types of operant conditioning, giving an example of each.
- Require students to think about and apply topics from class, not just repeat their notes.
 - Bad: State the four elements in observational learning.
 - Good: Describe how observational learning could be achieved in learning how to make pounded yam. Be sure to include a description of each of the four elements of observational learning.

Essay Items

- **Rubric:** Scoring scale that describes criteria for grading
- **Constructing Rubrics**
 - Establish criteria based on critical elements of essay
 - Determine levels of achievement for each criterion
 - Develop a clear description for each level of achievement
 - Determine rating scale for whole rubric

Example Rubric

Levels of Achievement			
Criteria	1	2	3
Topic Sentence	Not present; reader has no idea of what paragraph is about	Present but does not give the reader a clear idea of what the paragraph is about	Provides a clearly stated overview of the paragraph
Supporting Sentences	Rambling and unrelated to topic sentence	Provide additional information but not all focused on topic sentence	Provide supporting detail relating to topic sentence
Summarizing Sentence	Nonexistent or unrelated to preceding sentence	Relates to topic sentence but doesn't summarize information in paragraph	Accurately summarizes information in paragraph and is related to topic sentence

Example Marking Scheme

- a) Explain the four factors required for a multicultural society. **1.5 points apiece. TOTAL: 6**
- b) For each factor listed in part (a), describe whether this condition is currently met within the Jos environs and explain why. **1.5 points apiece. TOTAL: 6**
- c) Describe the concept of universal orientation as it relates to intercultural relationships. Be sure to state whether the universal orientation is beneficial or not. **2 points for definition, 4 for explanation. TOTAL 6**
- d) What can you as a teacher do to develop a universal orientation? Be specific. **TOTAL: 6 points**

Marking Essay Items

- Write a model answer for each item when setting the questions
- Prepare rubric or marking scheme when setting the questions
- Require students to answer each item
- Score all answers to one item before marking the next item
- Score all responses to single item in one sitting
- Score answers without knowing the identity of the student

Classroom Assessment

- Christine is just not very bright. She tries hard, but just doesn't have much cognitive ability.
- Monday is bright, but he is lazy. He knows when he has done enough and quits working.
- Mercy is bright, but she comes from a home in which English is not spoken. She will most likely do much better when she is proficient in English.
- John is a huge discipline problem and is frequently in school discipline for misbehavior.

Classroom Assessment

- What should grades measure?
 - Achievement in the specific class
 - Effort
 - Ability
 - Personality
 - Interests
 - Attitude

Sharing Assessment Outcomes with Parents

- Share child's academic, social, and character strengths and progresses
 - Share areas of concern with examples to illustrate
 - Collaborate to find strategies to help the child overcome areas of concern

Process of Developing a Formal Assessment

- Test Conceptualization: Idea for a test
- Test Construction: Drafting test items
- Test Tryout
- Item Analysis
- Test Revision

Test Conceptualization

- Step 1: Describe purpose and rationale for test
 - What the test will measure and for what purpose will the test will be used?
- Step 2: Describe the target population for the test
- Step 3: Clearly define the key variable of interest
- Step 4: Create item specifications
- Step 5: Choose item format
- Step 6: Specify administration and scoring procedures

Test Conceptualization: Purpose and Rationale

- What is the test designed to measure?
- What is the purpose for the test?
- Who will use the test?
- Who will take the test?
- How will the test be administered?
- What is the ideal format for the test given the test taker characteristics and time/financial considerations?
- Is there any potential harm that can result from test administration?
- How will test scores be meaningful?

Test Conceptualization: Target Population

- Relevant characteristics to consider:
 - Age
 - Educational Status
 - Language (which language and level of language ability)
 - Literacy Level
 - Disabilities

Test Conceptualization: Define Variable

- Develop by referencing:
 - Theory
 - Developmental guidelines
 - Empirical Literature
 - Cultural Definitions

Test Conceptualization: Item Specifications

- List of major content areas to be included in the test
 - Also can include number of items
- Developmental milestones to be achieved
- Otherwise, develop item specifications from theory or definitions of variable

Test Conceptualization Types of Instruments

- **Self-Report:** Participants report their own demographic characteristics, attitudes, beliefs, knowledge, feelings, and behavior
 - Can take the form of Questionnaire or Interview
- **Performance Assessment:** Directly assess performance on a contrived task
- **Observation Schedule:** Researchers observe participants' behavior under specified conditions
- **Checklist:** Identify the frequency or presence of behaviors or characteristics
- **Examination/Test:** Test participants' knowledge of a topic

Self-Report

- **Self-Report:** Participants are asked to report their own attitudes, beliefs, knowledge, feelings, and behavior
 - Self-report can be either a questionnaire or interview
 - Attitudes, beliefs, values, interests, knowledge, feelings, and some types of behavior are typically best assessed by self-report for most target populations

Questionnaire

- Advantages:
 - Easy to administer and score
 - Can be administered to larger numbers of test-takers
- Disadvantages:
 - The data is typically not as in-depth as an interview
 - Requires a high level of literacy

Interview

- Advantages:
 - Gathers rich, in-depth information
 - Can ask test-taker to expand on unclear responses
 - Can be used with any test-taker (e.g., cannot read)
- Disadvantages:
 - Takes time
 - Typically subject to issues of inter-rater reliability

Types of Items for Self-Report

- **Open-ended:** Test-takers give their own response
- Advantages:
 - More in-depth information
- Disadvantages:
 - Test-takers need to have good language/literacy skills
 - Test users have to categorize and code responses (scoring is subjective)
 - Subject to inter-rater reliability issues
 - Test-takers' responses may not be understandable (Questionnaires)
 - Test-takers often skip these items on questionnaires

Types of Items for Self-Report

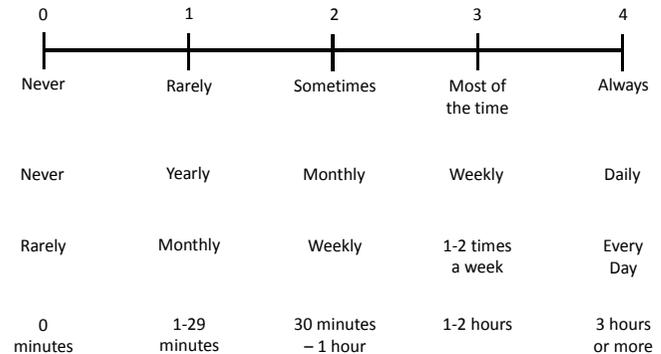
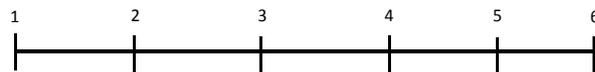
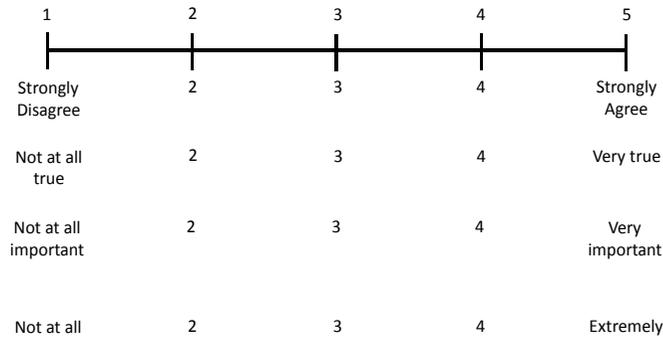
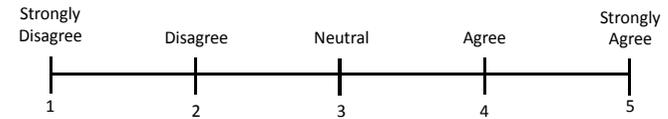
- **Close-ended:** Test-takers select their response from given categories on the test
- Advantages:
 - Easy and objective scoring
 - Response is clear to the test user
- Disadvantages:
 - Test developer must anticipate most common responses
 - Test developer must develop response categories that the target population understands
 - Test developer must develop response categories that are **exhaustive** (all participants have a place to appropriately respond)

Sample Items

- **Open-Ended:** *What causes you anxiety?*
- **Close-Ended:** *What causes you anxiety?*
 - a) Public speaking
 - b) Large crowds
 - c) Examinations
 - d) Driving

Test Item Formats

- **Rating Scale:** Grouping of words, statements, or symbols on which judgments concerning the strength of a particular trait, attitude, or emotion are indicated by the test-taker.



Test Format: Checklists

- Checklists identify the frequency or presence of behaviors or characteristics
 - Trauma Symptom Checklist for Young Children

Your Baby at 2 Months

Child's Name _____ Child's Age _____ Today's Date _____

How your child plays, learns, speaks, acts, and moves offers important clues about your child's development. Developmental milestones are things most children can do by a certain age. Check the milestones your child has reached by the end of 2 months. Take this with you and talk with your child's doctor at every visit about the milestones your child has reached and what to expect next.

What Most Babies Do at this Age:

Social/Emotional

- Begins to smile at people
- Can briefly calm himself (may bring hands to mouth and suck on hand)
- Tries to look at parent

Language/Communication

- Coos, makes gurgling sounds
- Turns head toward sounds

Cognitive (learning, thinking, problem-solving)

- Pays attention to faces
- Begins to follow things with eyes and recognize people at a distance
- Begins to act bored (cries, fussy) if activity doesn't change

Movement/Physical Development

- Can hold head up and begins to push up when lying on tummy
- Makes smoother movements with arms and legs

Act Early by Talking to Your Child's Doctor if Your Child:

- Doesn't respond to loud sounds
- Doesn't watch things as they move
- Doesn't smile at people
- Doesn't bring hands to mouth
- Can't hold head up when pushing up when on tummy

Tell your child's doctor or nurse if you notice any of these signs of possible developmental delay for this age, and talk with someone in your community who is familiar with services for young children in your area, such as your state's public early intervention program. For more information, go to www.cdc.gov/concerned or call 1-800-CDC-INFO (1-800-232-4636).

Adapted from CARES FOR YOUR BABY AND YOUNG CHILD: BIRTH TO AGE 5. Film Edition, written by Steven Shaker and Tanya Renner Ahmann © 1991, 1993, 1998, 2004, 2009 by the American Academy of Pediatrics and SPECIAL FEATURES: GUIDELINES FOR HEALTH SUPERVISORS OF INFANTS, CHILDREN, AND ADOLESCENTS, Third Edition, edited by Joseph Hagan, Jr., Judith S. Shaw, and Paula M. Dawson, 2006. © American Academy of Pediatrics. This milestone checklist is not a substitute for a standardized, validated developmental screening tool.

Test Format: Performance Assessment

- Directly assess behavioral performance on tasks
 - Self-regulation
 - Creativity
 - Social skills
 - Identification of letters
 - Gross motor skills

Test Format: Observation

- Measures behaviors only
- Examples
 - Parent-Child Interaction Assessment-II (PCIA)
 - Parent-Child Early Relational Assessment
 - Dyadic Parent-Child Interaction Coding System-II



Test Format: Observation

- Consider:
 - Where and when will the behaviors be observed?
 - What possible situation will the test-taker will be exposed to?
 - How long will test-takers be observed?
 - How will behaviors be recorded?
 - **Duration:** Length of time that behaviors occurs
 - **Frequency Count**
 - **Interval:** Set a stopwatch for a specific interval of time, then record what behavior the test-taker is engaged in

Item Construction Developing Good Items

- Use clear words so all participants understand
- Be precise and specific in the wording of each item so different test-takers do not interpret the item differently
- Ensure the item is not biased or leading toward one response.
- Each item should have only one point (avoid double-barreled items)
- Ensure that each test-taker has a place to respond

Administration and Scoring Guidelines

- Standardized assessment requires very clear administration guidelines so all participants complete the assessment under similar conditions
 - Critical for performance and observation assessments
- Scoring guidelines should provide specific score interpretations based on empirical research

Test Try-Out

- Try items out on people similar in critical respects to the people for whom the test was designed
- Sample size: No fewer than 5 participants (preferably at least 10 participants) for every 1 item on the test
- Conduct under conditions as identical as possible to conditions under which the final test will be administered