Assessment Literacy: From Theory to Practice

Webinar Series
February-April 2009

Facilitated by Deb Farrington

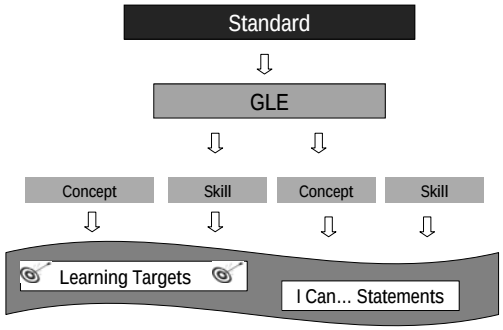
Agenda

- Evaluating Depth of Knowledge
- Assessment Design Considerations
  -Multiple Choice Items
  -Constructed Response Tasks

What Constitutes Effective Classroom Assessment?

- provides evidence of student performance relative to content and performance standards
- provides teachers and students with insight into student errors and misunderstanding
- helps lead the teacher directly to action

The Dichotomy of an Unpacked Standard



Why Depth of Knowledge?

- Mechanism to ensure that the intent of the standard and the level of student demonstration required by that standard matches the assessment items
- Provides cognitive processing ceiling (highest level students can be assessed) for item development
- Descriptive, not a taxonomy
- Not the same as difficulty

Webb's Depth of Knowledge Levels

Recall and Reproduction:	Level 1
Skills & Concepts:	Level 2
Strategic Thinking:	Level 3
Extended Thinking:	Level 4

Recall and Reproduction: Level 1

DOK 1 requires recall of information, such as a fact, definition, term, or performance of a simple process or procedure.

Answering a Level 1 item can involve following a simple, well-known procedure or formula. Simple skills and abilities or recall characterize DOK 1.

Skills/Concepts: Level 2

DOK 2 includes the engagement of some mental processing beyond recalling or reproducing a response. Items require students to make some decisions as to how to approach the question or problem.

•These actions imply more than one mental or cognitive process/step.

Strategic Thinking: Level 3

DOK 3 requires deep understanding as exhibited through planning, using evidence, and more demanding cognitive reasoning. The cognitive demands at Level 3 are complex and abstract.

An assessment item that has more than one possible answer and requires students to justify the response they give would most likely be a Level 3.

Extended Thinking- Level 4

DOK 4 requires high cognitive demand and is very complex. Students are expected to make connections—relate ideas *within* the content or *among* content areas—and have to select or devise one approach among many alternatives on how the situation can be solved.

Due to the complexity of cognitive demand, DOK 4 often requires an extended period of time.

However, extended time alone is not the distinguishing factor.

Task	*Thinking*
Collecting data samples over several months	Recall
Organizing the data in a chart	Skills/ concepts
Using this chart to make and justify predictions	Strategic Thinking
Developing a generalized model from this data and applying it to a new situation	Extending Thinking

The Depth of Knowledge is NOT determined by the verb, but the context in which the verb is used and the depth of thinking required.

Same verb—three DOK levels

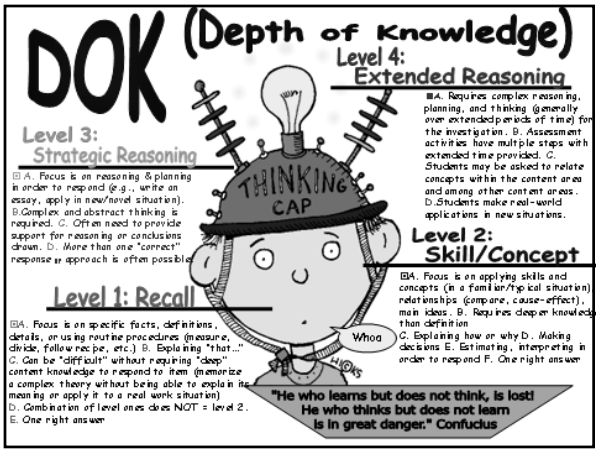
- DOK 3- ***Describe*** a model that you might use to represent the relationships that exist within the rock cycle. (requires deep understanding of rock cycle and a determination of how best to represent it)
- DOK 2- ***Describe*** the difference between metamorphic and igneous rocks. (requires cognitive processing to determine the differences in the two rock types)
- DOK 1- ***Describe*** three characteristics of metamorphic rocks. (simple recall)

Remember…

Depth of Knowledge (DOK) is a scale of cognitive demand. A cognitive ceiling.

• DOK requires looking at the assessment item/standard-not student work-in order to determine the level. DOK is about the item/standard-not the student.

• The context of the assessment item/standard must be considered to determine the DOK-not just a look at what verb was chosen.



Developing Classroom Assessment Items

- General item development tips
- Developing multiple-choice items
- Developing constructed-response tasks

General Item-Writing Guidelines

Use simple, basic vocabulary instead of technical vocabulary unless you are assessing the students' knowledge of the *meaning* of the technical word/phrase.

General Item-Writing Guidelines

- Use grade-appropriate vocabulary as much as possible
- Eliminate irrelevant information
- Include clear, correct, easily understood graphics as **required** by the item
- Consider ***Bias and Sensitivity Guidelines*** when drafting items

Avoiding Bias

Items should focus on academic experience rather than life experience throughout

- question/writing prompts
- identified situations
- graphics
- reading selections

Advantages of Multiple-Choice Items

- Can be used to measure a wide variety of learning outcomes
- Permit wide sampling and broad coverage of a content domain
- Are reliable and efficient to score
- Can provide useful diagnostic information about the learning of individual students or groups of students

Disadvantages of Multiple-Choice Items

- Multiple-choice items can be difficult to write to certain content or DOK targets
- Multiple-choice items are not well suited for measuring certain types of skills (e.g., the ability to organize and express ideas in writing).
- Performance on multiple-choice items can be influenced by student characteristics unrelated to the subject of measurement, such as reading ability and “test wiseness.”

Multiple-Choice Item Guidelines

- Be sure there is only one right answer
- Present a single, clear, complete problem or question in the stem of the item
- State the item stems in positive terms (if possible)
- Avoid using negatives in both the item stem and the answer choices; double negatives are confusing

Multiple-Choice Item Guidelines

- Avoid the use of absolute terms in answer choices (always, never, all, none, only)
- Whenever possible, avoid answer choices that are mutually exclusive opposites (e.g., fiction/nonfiction, living/non-living). When such opposites are used, a student's chance of getting the item correct becomes 1 in 2 versus 1 in 4.

Response Options (is it easy as ABC…D?)

- Be sure the correct answer (key) is indisputable
- Make the incorrect response options (distractors) plausible
- Develop parallel response options; options should be parallel with respect to content, structure and length

Response Option "Rules"

- Common misconceptions or errors of students are good response options to include in an item
- Avoid using the options "all of the above" and "none of the above"
- Avoid humorous or nonsensical response options
- Logically order the options

Advantages of Constructed-Response Items

- Constructed-response or open- response items allow for deeper depth of knowledge to be demonstrated than in multiple-choice items.
- Students can be asked to demonstrate more complex cognitive behaviors such as comparing, relating, analyzing, inferring, concluding, predicting, generalizing, solving and/or applying.

Disadvantages of Constructed-Response Items

- Constructed-response items are more difficult and more expensive to score.
- Because of their cost, the relative number of constructed-response items is significantly less than multiple choice.
- Effectiveness of constructed-response items is based on the scoring guide and answer information provided.

Make sure that the task is actually achievable

If the item asks students to read a passage or examine a graphic and then "give three ways" or "explain two reasons" based on the material given, make sure there are three ways or two reasons found in the material provided.

Make sure the directions are complete and specify what is desired from the students

- If you want the students to provide examples, the directions to the students should tell them to provide examples and how many.
- If you want the students to identify information, do not ask them to discuss, describe or explain.

Make sure that each item has an item prompt that precedes the directions

Even if the item has a graphic, there still must be a prompt that describes or provides information related to the graphic and/or item directions.

Suggestions for Item Format

- Present the prompt in paragraph form.
- Use bullets to emphasize the details in the prompt.
- If the students are required to respond to multiple parts of a question, label each part separately (a, b, c).

Sample Constructed- Response Item

Martin said, "I am thinking of a whole number between 100 and 300. The number is divisible by three but not by 9. The ones digit is the sum of the hundreds digit and the tens digit."

a. Show why 153 cannot be Martin's number.
b. Find **all** the numbers that match Martin's clues. Show all your work.
c. Write one more clue that would limit the answer in **part b** to one and only one correct number.

Action Words Found in GLE's

Describe	Explain	Identify
Illustrate	Trace	Compare
Contrast	Predict	Apply
Sort	Justify	Analyze
Evaluate	Discuss	Define
List	Differentiate	Distinguish
Defend	Conclude	Categorize

Openness

- Questions that are too *open* are difficult for students to manage, and may provide little evidence of the learning to be examined
- Questions that are too *closed* may not provide enough breadth of performance, and may be better suited to another question type

Scorability

- Does the question provide for distinct levels of performance that are scorable?
- Are distinctions between score points important or trivial?

Designing Classroom Assessments
Fundamental Questions

- What learning target are you assessing?
- What depth of knowledge is expected?
- What kind of evidence would you like to gather?
- What is the best assessment method for gathering this information?
- How will performance be assessed?