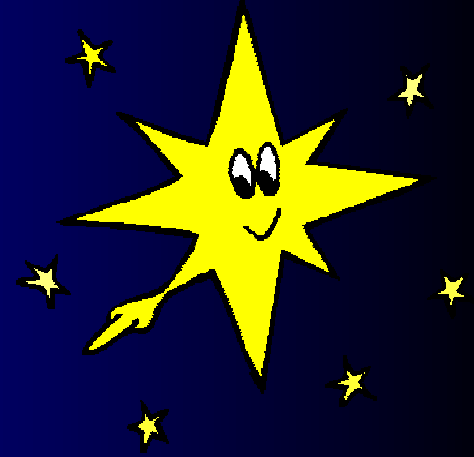


# Instructionally Oriented Assessments



By: Gary Whiteley

# Utilizing Instructionally-oriented Assessments in the Classroom

- Content and Construct Validity
- Review and create item-types
- Link specific GLE's to assessments
- Reading in the middle and high school

# Activity: With an Elbow Partner for 10 minutes

Recall the research finding:

*The impact of improved classroom assessment is greater than that of most educational interventions.*

1. From the perspective of a classroom teacher, which two of the five essential questions are the most challenging?
2. Needed: A moderator to keep time and close-out the activity.

# Essential questions for teachers:

- What do I teach? (develop assessments first)
- How long do I teach it? (case-by-case)
- What should I assess? (sampling)
- How should I assess it? (item-types)
- Was the instruction effective? (student scores)
- Do the students have specific feedback in order to improve? (evidence v. opinion)

# Validity: It Really Matters: What to Assess and How to Assess

## Content Validity

The process used for developing the Grade Level Expectations (GLE's) took into account the relationship between the Alaska Content and Performance Standards and each individual GLE. Does the item measure the knowledge or skill?



*(The GLE's are not intended to be the curriculum. The GLE's should be "covered" by a well-developed curriculum. They are statements of essential skills developed by Alaskan educators.)*

## Construct Validity

Construct validity is often overlooked when developing classroom assessments. Which item-type should we use to measure a "construct"?

Assessments do not possess validity (score-based inference)

# Item-types:

## A Deliberate Choice

*There is no “best” item type. The type of information assessed, the time available for administering and scoring, the scoring criteria, and the depth of student knowledge required on the assessment all factor into the selection of an item type.*

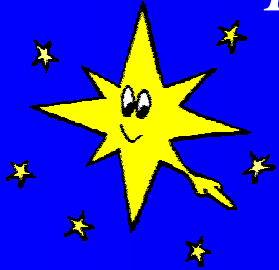
# Choices, Choices, Choices

- 1. Short-Answer Items**
- 2. Matching Items**
- 3. Binary Choice Items**
- 4. Multiple Binary Choice**
- 5. Multiple Choice Items**
- 6. Essay Items**
- 7. Performance Task Items**

Source: *Classroom Assessment: What Teacher Need to Know*. (J. Popham)

# 1. Short-Answer Items

*Definition: A student generated written response to an item.*

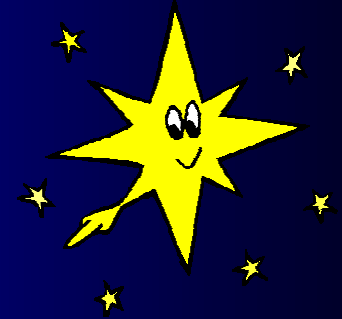


- Verbal (word, phrase or sentence)
- Not Figurative (no diagrams or pictures)
- Shorter than an essay

# How to achieve validity?

- Define criteria for scoring within instructions
- Use direct questions vs. incomplete statements
- Construct the item so the answer will be concise
- Blank placement is important
- Put blanks at the end of incomplete statements.
- Limit answer space or blanks to foster "concision"
- Keep all lines the same length

## 2. Matching Items



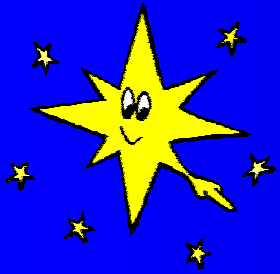
- Employ homogeneous lists  
*questions and answers are parallel, e.g.,*
- Definitions and vocabulary  
*Presidents with their challenges*
- Use brief lists  
*between 5 to 10, shorter words on the right*
- Use more responses than premises  
*add extra, making sure there are no unintentional clues*

# How to achieve validity?

- Describe basis for matching
- Place on single page
- Order logically, list chronological or alphabetically
- Describe the relationship you want the student to find
- Matching is not appropriate for distinct ideas
- All matching items must be related

# 3. Binary Choice

*Definition: A student generated selection between two listed items*

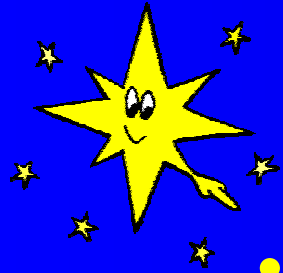


- True / False
- Yes / No
- Correct / Incorrect
- Fact / Opinion

# How to achieve validity?

- Phrase items so a superficial analysis suggests a wrong answer
- Rarely use negative statements
- Never use double negatives
- Include only one concept in each statement
- Have an approximately equal number of items representing the two categories being tested
- Keep item length similar for both categories being tested

# 4. Multiple Binary Choices



*How does it work?*

- Statement- stem/stimulus is presented as the overlying base for the questions to follow
- Then binary response items follow
  - True / False
  - Yes / No
  - Correct / Incorrect
  - Fact / Opinion

# Advantages to using Binary-choice items

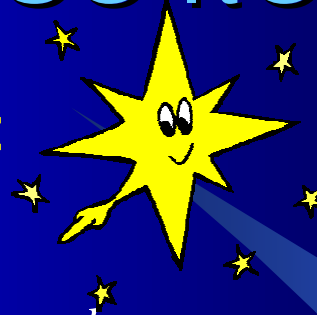
- Higher level of Bloom's taxonomy when new content in the stem is introduced
- Highly efficient for gathering student data
- More reliable than other selected response items for measuring understanding
- More difficult for students and more efficient for teachers than multiple choice

# How to achieve validity?

- Separate item clusters
- Each item meshes well
- Similar length of questions
- Questions stay on topic with the stem
- Do not use negative statements

# 5. Multiple Choice Items

## Parts of a multiple choice item:



- **Stem**

*The core of a multiple-choice item, stated as a stand-alone question or statement*

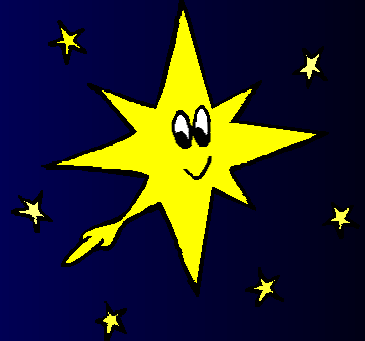
- **Item Alternatives**

*The choices offered as responses to the stem. It is recommended that 4 or 5 alternatives be given. Never less than 4, rarely more than 4*

- **Distractors**

*Incorrect alternatives-Each distractor should be plausible and should relate clearly to the stem*

# Multiple Choice: Guidelines for Writing



- The stem should consist of a self-contained question or problem
- The item should not rely on the alternatives in order to make sense, instead the stem should lead to a correct response without the alternatives listed
- Avoid negatively stated stems
  - We all dreaded the questions using the word “not”.*
- Do not allow the length of alternatives to supply unintended clues. Alternatives that are unusually lengthy or short in comparison to the other alternatives may cause students to focus on these alternatives as being correct when, in fact they may not be

# Multiple Choice: Guidelines for Writing Cont.

- Randomly assign correct answers to alternative positions
- Alternatives A, B, C, and D should be listed as correct responses an equal number of times through out an assessment
- Never use “all-of-the-above” alternatives, but “none-of-the-above” alternatives may be used

*An “all-of-the-above” alternative may be overlooked when students encounter a correct response in the first alternative. Additionally, if a student only knows that two of the alternatives are correct but is unsure about the third, s/he will know to select “all-of-the-above”. “None-of-the-above” does not present these problems and may offer an increased level of difficulty when desired.*

## 6. Essay Items/Constructed Response



- Convey to students a clear idea regarding the extensiveness of the response desired
- Construct items so that the student's task is explicitly described
- Provide students with the approximate time to be expended on each item as well as the item's value
- Precursively judge an item's quality by composing (mentally or in writing) a possible response
- Avoid optionality in essay items

# Essay Items: Guidelines for Scoring

- Scores items holistically and/or analytically
- Prepare a tentative scoring key in advance of judging student responses
- Make decisions regarding the importance of the mechanics of writing prior to scoring
- Score all responses to one item before scoring responses to the next item
- Insofar as possible, evaluate responses anonymously

# 7. Performance Task Items

- **Make sure the student's performance at the task will generalize to comparable tasks.**
- **Make sure the task is realistically possible in terms of cost, space, time, and equipment needed.**
- **Make the task something the student might encounter in the real world, not just at school.**
- **Make the task measure multiple skills, not just one.**

# Performance Task Items

- **Make the task one that the student can become more proficient at as a consequence of instruction.**
- **Make the task fair to all students; avoid bias.**
- **Make the task such that student response can be reliably and accurately measured.**
- **Make students aware in advance of the criteria upon which their work will be judged.**

# Rubrics : Areas of Concern

## Scales that do not differentiate

*Lack of descriptive language between cells creates challenges in making distinctions*

## Lack of descriptive rigor

*Language that is too vague or too specific*

## Scales that are too easy or too difficult

*This is a major concern when developing a rubric that has high stakes implications*

Teachers on the committee compiled rubrics for you. (Three point scale)

(Reliability studies and Exemplars needed)

A few words about Readability.

# We selected Reading GLE's to look at because...

*Seventy-five percent of students with literacy problems in the third grade will still experience literacy difficulties in the ninth grade.*

*Approximately 25 percent of all high school students read below basic levels or three to four years below basic grade levels.*

*The National Center for Education Statistics (NCES) reports that Approximately 45–50 percent of African American, Hispanic, and American Indian students score below the basic level on the NAEP.*

The information on the previous slide is from  
**Creating a Culture of Literacy: A Guide  
for Middle and High School Principals**

- Download the report at the National Association of Secondary School Principals web site: [www.principals.org](http://www.principals.org)
- Since our topic is assessment, a superb list of diagnostic and progress-monitoring assessments can be found in:

Appendix 5: Assessments for Middle and High School Students

# Training Activity



1. A table moderator who will keep time and help close-out the activity

2. You will need the sheets of paper with the header **“Item Writing Guidelines Summary and Essay Item Writing/Scoring”**



# Elbow Partner Activity

**Total time: 20 minutes (5 min. for each task)**

**Moderator please keep time and remind others**

1. Turn to pages 54-59 in standards book to review GLEs within the “Analysis of Content or Structure” strand
2. Now turn to page 54 and read -Grade 6- The student analyzes content and structure of genres by: (2.7.1), (2.7.2), (2.7.3.) identify or explain
3. Discuss with your partner(s) how you would assess this GLE 2.7.3. Including the item-type(s) you would use and reason you chose those items
4. Create one item-type.

*Only 674 left to go!*

## 2.7.3 (6<sup>th</sup> grade) Identifying or explaining use of literary elements and devices appropriate to genre

(i.e.: dialogue, rhyme, alliteration, simile, metaphor or personification)

**Directions: Match the literary element or device to its correct definition.**

- a) gives real-life characteristics to non-living objects
- b) a comparison made between two things using the words like “or” and “as”
- c) similarity in the sounds of words
- d) words spoken by characters in a book
- e) figure of speech in which a word or phrase meaning one kind of object or idea is used in place of another to suggest a similarity between them
- f) literary effect achieved by using several words that begin with the same or a similar consonant

The following choices appear to the left of the stems: Alliteration, Dialogue, Personification, Metaphor, Simile

# Elbow Partner Activity

**Total time: 14 minutes (7 min. for each task)**

1. Now turn to page 55 and read -Grade 10-  
The student analyzes and evaluates  
conventions and techniques of genres by:  
4.5.1, 4,5.2, 4.5.3.
2. Discuss with your partner(s) how you  
would assess this GLE 4.5.2, including the  
item-type(s) you would use and why you  
think it is a good choice.

## 4.5.2 (10th Grade) Analyzing the use of literary devices appropriate to genre

(i.e. dialogue, simile, metaphor, personification, foreshadowing, time sequence, imagery, or repetition, allusion or symbolism) to analyze literary works.

**Directions: Identify the type of each literary device used below. Use the term bank to assist in completing this exercise.**

TERM BANK	EXAMPLE	DEVICE
Simile Personification Repetition Foreshadowing Time Sequence Dialogue Metaphor Imagery		

It may be useful to determine the frequency that the GLE's are covered in a school year.

## GLE Audit Juneau School District

**R. 2.1 Reading GLE's grade 6**  
The student uses strategies to  
Decode or comprehend meaning  
Of words in text by:

To what degree is the GLE covered?

0 \_\_\_\_\_

1-3 \_\_\_\_\_

4-6 \_\_\_\_\_

7-10 \_\_\_\_\_

**2.11 Demonstrating knowledge of  
word structure...**

**Evidence:**

**Where is the GLE addressed in the curriculum/CORE?  
List units/projects, and other venues where the GLE's  
are addressed in your content area.**

Thank you  
Janet Lopez!

# Table Activity for 10 minutes

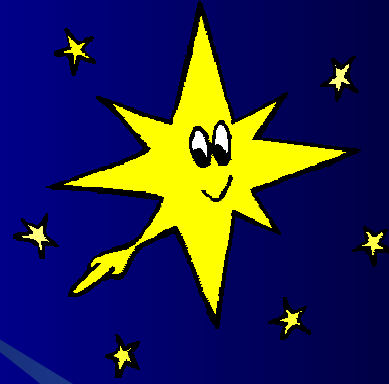
Needed: Moderator to time and facilitate

Please locate the self-assessment that you completed earlier today.

In light of the information presented and discussions with your colleagues:

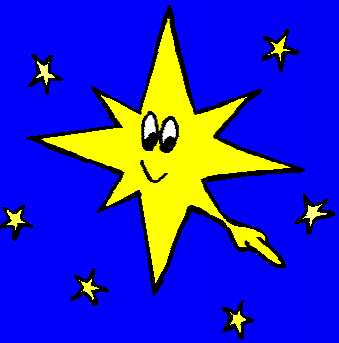
Are there any responses that you would change?

# Final Thoughts



- Validity! It Really Matters
- Content Validity - The GLE's!
- Construct Validity
- There is no “best” item type
- But there are poorly designed items and items that do not assess the content.

- Designed by **teachers** for **teachers** with **teaching** and **assessing student learning** as the focus
- The Alaska Assessment Team brings **over 800 years of teaching experience**
- The Alaska Assessments are not perfect, however our combined efforts created “**capacity**” we could not have accomplished alone
- Coming Soon, **available in mid-February**



# The Role of the District: Remember the research

- Professional Development will be necessary.
- A process with a protocol to examine, analyze, and discuss the data. Develop exemplars!
- Consider backward planning for the use of the assessments.
- Continue the discussion by using the self-assessment, assessment types, and essential questions.

# And remember...

*“Change is for the person on the ladder rung below you.”*

*-Roland Barth*

School Leadership is a must!

Please include teachers in the definition of school leadership.