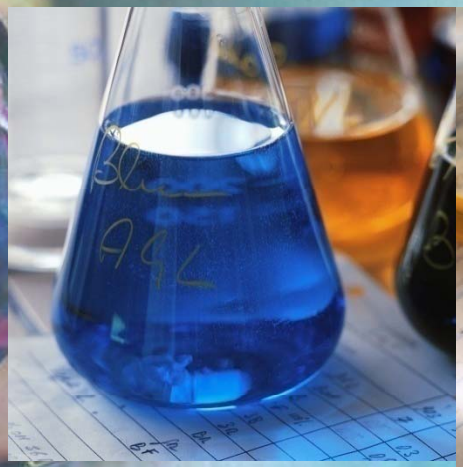
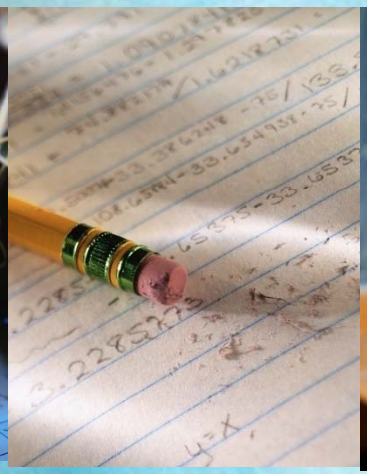
Miami-Dade County Public Schools

Best Practices in Constructing Assessments

Benchmark Assessment Item Bank





Differentiated Accountability Training

December 3-8, 2008

What is the item bank?

- Searchable electronic database of test items that allows teachers to create tests
- High quality items that are aligned to the benchmarks/standards that are currently being taught
- Items that differ in complexity and type (multiple choice, short, extended, and gridded responses)
- Allows the teacher to retrieve test items based on specific needs

Purpose of the item bank?

- **To allow teachers to monitor student progress on the Reading, Mathematics, and Science Sunshine State Standards**
- **To provide teachers with classroom assessment tools that will provide reliable student-level benchmark information to inform instruction**
- **To provide students with information on their progress on specific benchmarks**

Number of Items

Grade Level	Math	Reading	Science
3	535	269	N/A
4	532	250	345
5	556	245	325
6	530	289	N/A
7	599	268	451
8	525	281	407
9	469	283	N/A
10	537	280	360
11	N/A	N/A	380

Steps to Constructing a Classroom Test

- ❑ Determine the purpose of the test
- ❑ Construct a test blueprint
  - -Benchmarks to be assessed
  - -Cognitive complexity
  - -Length of test/number of items
- ❑ Types of items that will be used
- ❑ Select appropriate and relevant test items
- ❑ Assemble the test
- ❑ Administer the test
- ❑ Use the results

Purpose of the Test

Determine intended use of the assessment:

- Instructional decisions – spend more time on a skill, skip the review, or stick to the instructional plan
- Grading decisions- end of unit test, midterm, or final exam
- Diagnostic decisions- determine students' strengths and weaknesses
- Remediation decisions

Getting Started: Developing A Blueprint

- Identify standards/benchmarks to be tested
- Percent of items by content focus
- List targeted skills that you want assessed
- Specify content focus
- Number of items
- Cognitive level

Selecting Types of Test Items

- Item provides a direct measure of the outcome it was designed to measure
- Select the test-item type most appropriate for the purpose
- Analyze the performance and depth described in the specific learning outcome
- Assess the depth of student knowledge
- Include a balance of item difficulty using Webb's Levels of Complexity

Building the Test Blueprint

- Design a **Test Blueprint (or Table of Specifications)** for the assessment that you will build using ExamView

Standard	Webb's Depth of Knowledge			Total Number of Items
Benchmark MA.D.1.2.1: The student describes a wide variety of patterns and relationships through models, such as manipulatives, tables, graphs, rules using algebraic symbols.	Low	Moderate	High	
Extend numeric pattern with list				
Describe numeric pattern with chart				
Extend geometric pattern				
Identify predictable pattern				
Total Number of items				
Percent of items				

Selecting Appropriate and Relevant Items

- Does the item measure the benchmark?
- Does the item measure an important concept?
- Is the item appropriate for the grade level?
- Does the item require students to use the appropriate cognitive level?
- Are there different levels of complexity reflected?
- Is the level of reading skill required by this item below the students' ability? (mathematics and science)
- Will all or almost all students answer the item in the suggested time frame for the type of question?

Assembling the Test

Decide length

- Minimum of four items for each, if you want a reliable measure of an individual benchmark
- Number of items: will depend on complexity, grade, and purpose of test
- Length of test administration should be appropriate for grade level

Assembling the test

- Check your Answer Key
- Determine how students will record their answers
- Provide space for name and date
- Be sure directions are specific and accurate
- Proofread the test before distributing it to students

Selecting Items From ExamView

The ExamView Test Generator allows you to create and customize the test with a variety of selections:

- Standard
- Criteria
- Viewing questions
- Randomly

Administering the Test

- Students should have enough time to complete all items
- Provide instructions before the test is administered
- Alert students to the amount of time left toward the end of the test
- Clarify test collection procedures before handing out the test
- Maintain a positive test-taking attitude

Scoring the test

- Have the key prepared in advance
- Check the key for accuracy
- Decide if you will hand score or use the Edusoft System to develop and score your answer sheets
- Ensure that you have given each item the appropriate point value
- Use the rubrics and anchor papers for scoring guidelines for constructed test items (short, and extended responses)

Evaluate the test

Select the report that is appropriate to the needs of the teacher.

Possible Benchmark Exam Reports that can be retrieved via Edusoft:

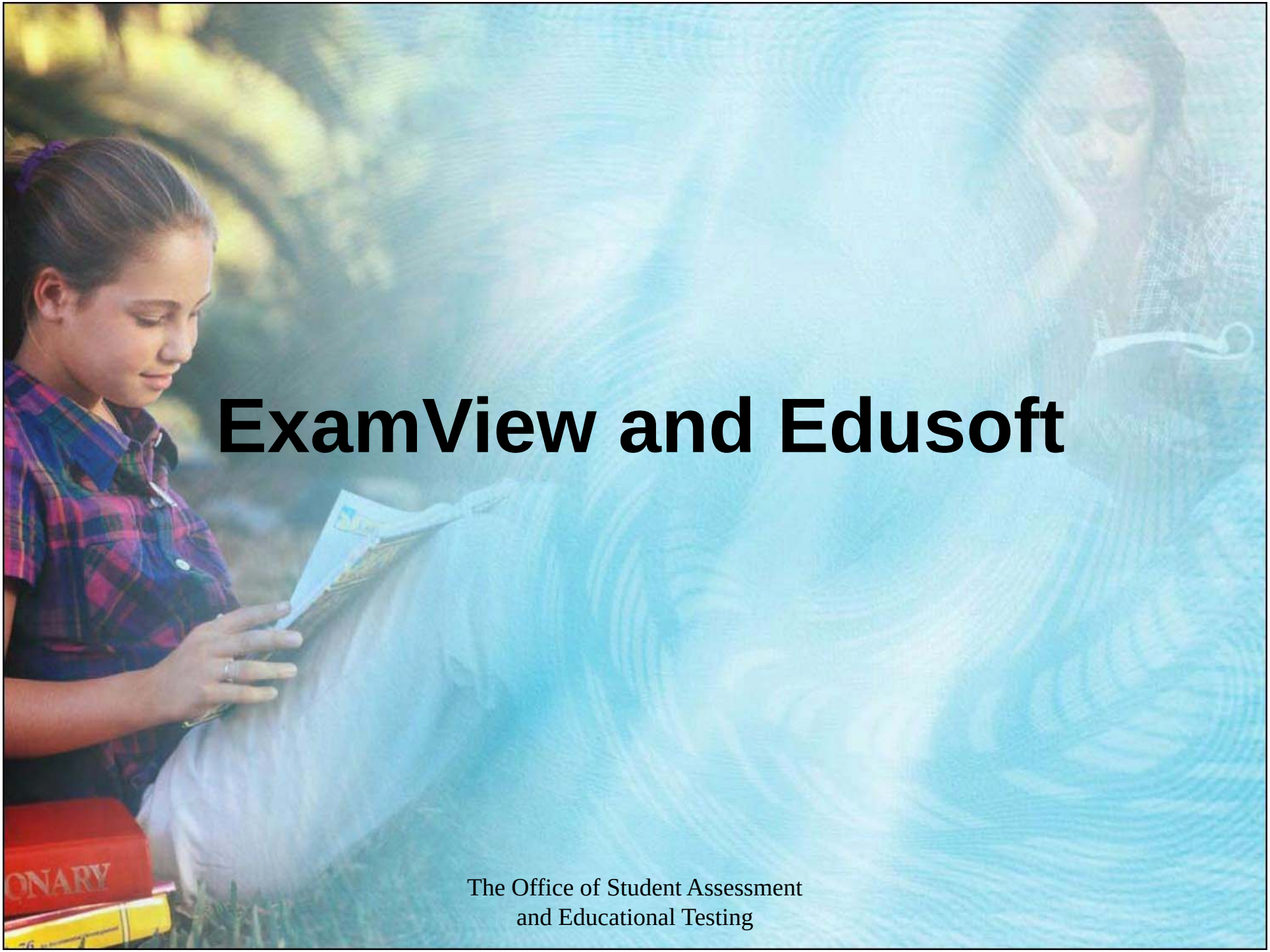
- ❑ Class List Reports
- ❑ Item Analysis Report
- ❑ Item Response Report
- ❑ Student Performance Report

Use the Results (Debriefing Guidelines)

- Provide students with their test and answer sheet
- Examine the test with your students
- Discuss any items you found to be problematic for the class as a whole
- Listen to students comments and reactions
- Make notes of any deficient or problematic items during this process
- Consider students comments, complaints, and suggestions for test improvement

Reviewing the Effectiveness of the Classroom Test

- Did the test accomplish what was intended?
- Were you able to use the test results to determine progress towards learning benchmarks?
- Were students able to use the test results to determine progress?
- What did students think about the test?



ExamView and Edusoft

The Office of Student Assessment
and Educational Testing

Creating a Classroom Assessment in Examview

Strand D: Algebraic Thinking
Standard 1: The student describes, analyzes, and generalizes a wide variety of patterns, relations, and functions.

Benchmark MA.D.1.2.1: The student describes a wide variety of patterns and relationships through models, such as manipulatives, tables, graphs, rules using algebraic symbols.

Fifth (What should the items look like?)
•1. describes, extends, creates, predicts, and generalizes numerical and geometric patterns using a variety of models (for example, lists, tables, graphs, charts, diagrams, calendar math).
•2. poses and solves problems by identifying a predictable visual or numerical pattern such as: Day 1 2 34 … *n* Number of Calls 4 7 10 ? ?
•3. explains and expresses numerical relationships and pattern generalizations, using algebraic symbols (for example, in the problem above, the number of calls on the *n*th day can be expressed as 3*n*+1).

Creating a Test in Examview

- Download the Examview generator from http://it.dadeschools.net/examview.htm to your desktop.

- Click twice on the icon on the desktop.

Interpreting Test Results: Item Analysis Report

The Kuder-Richardson 20 reliability value for this exam with these student selections is .88

What do these terms mean?

Analyzing 346 students

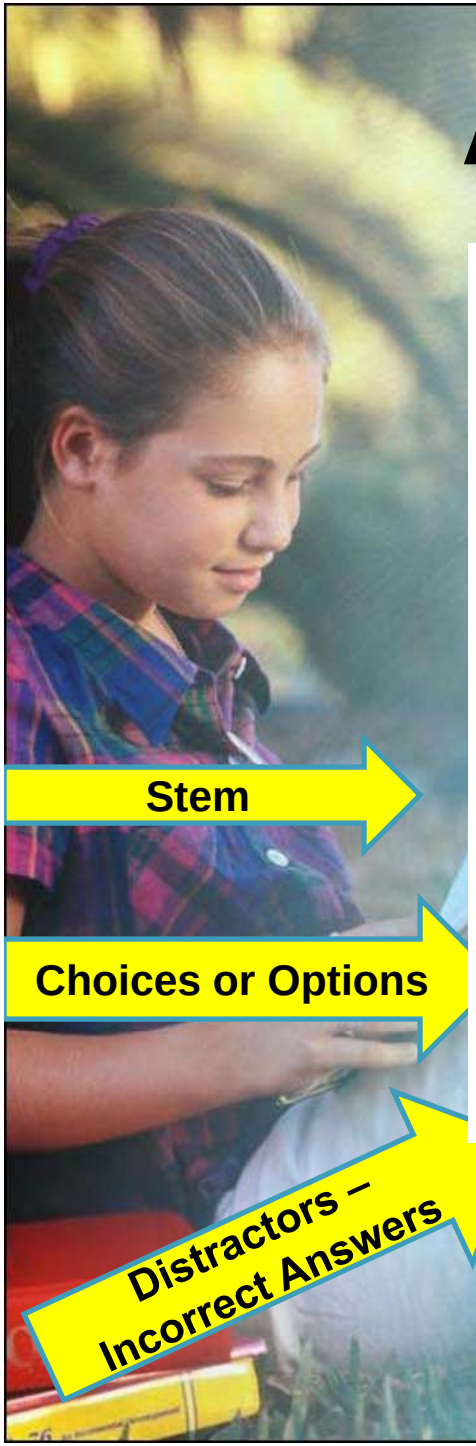
Group by: Section GO

Item Analysis - Multiple Choice questions

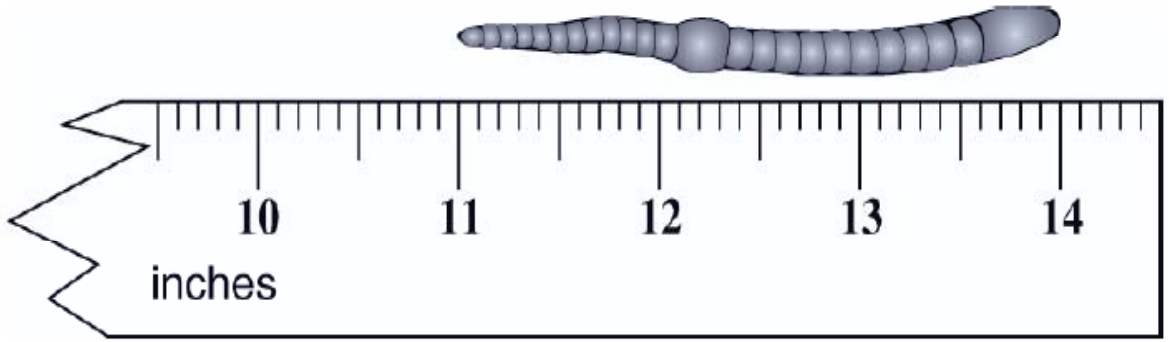
Raw Number | Percent | P-Biserial

Item	% Correct	Discrimination	A	B	C	D	F	G	H	I	OMITTED
sort by	sort by	sort by	sort by	sort by	sort by	sort by	sort by	sort by	sort by	sort by	sort by
Section 1 -> 1	78%	0.37	6.6	5.8	78.3 * –	9	0	0	0	0	0.3
Section 1 -> 2	79%	0.41	0	0	0	0	4.6	4.9	79.5 * –	11	0
Section 1 -> 3	86%	0.24	85.8 * –	5.2	3.5	5.2	0	0	0	0	0.3
Section 1 -> 4	93%	0.23	0	0	0	0	2.3	2.3	93.4 * –	2	0
Section 1 -> 5	84%	0.33	84.4 * –	3.5	4	8.1	0	0	0	0	0
Section 1 -> 6	54%	0.45	0	0	0	0	54*	17.6	5.8	22.5	0
Section 1 -> 7	40%	0.43	39.6 * –	9.5	24.3	26.6	0	0	0	0	0
Section 1 -> 8	55%	0.45	0	0	0	0	8.4	16.5	55.2 * –	19.7	0.3

Anatomy of a Test Item



29. Peter found a worm like the one shown below.



Which BEST represents the length of the worm?

A. 3 inches ★

B. 4 inches

C. 11 inches

D. 14 inches

Stem

Choices or Options

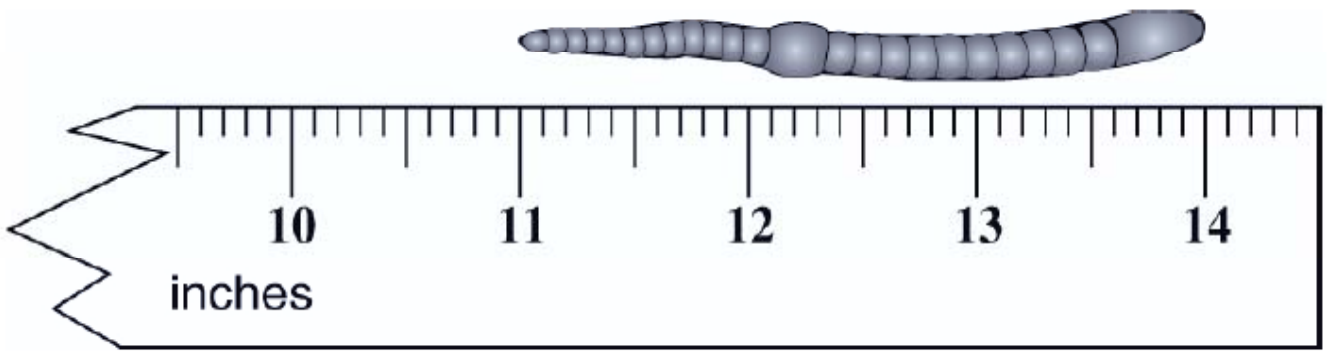
Distractors – Incorrect Answers

Key or Correct Answer
*** A is the correct answer**

Interpreting Student Responses

Item	% Correct	Discrimination	A	B	C	D	F	G	H	I	OMITTED
Multiple-Choice -> 29	11%	0.13	**10.9***	23.5	18.7	46.2	0	0	0	0	0.8

29. Peter found a worm like the one shown below.



Which BEST represents the length of the worm?

A. 3 inches

B. 4 inches

C. 11 inches

D. 14 inches

Interpreting Test Results: Item Response Report

# of Students: 18							
		Question #	1	2	3	4	5
		% of Correct Responses	89%	89%	39%	78%	33%
		Total Points Possible	1	1	1	1	1
		Correct Response	C	H	A	F	C
Student Name	Raw Score	% Correct			-		
Sebastien	21	64%	√	√	C	G	A
Jonathan	11	33%	A	F	√	I	A
Samantha	20	61%	√	√	D	√	A
Carlos	14	42%	√	√	√	√	D
Brianna	18	55%	√	√	C	√	A
Destinee	15	45%	√	√	D	√	√
Jason	17	52%	√	√	C	√	A
Breyana	20	61%	√	√	√	√	√
Matthew	13	39%	√	√	√	√	B
Janice	19	58%	√	√	√	√	B
Isaac	23	70%	√	√	D	√	√
Cedrick	5	15%	√	G	C	G	A
Devitch	12	36%	√	√	C	G	A
Daniel	8	24%	√	√	B	√	B
Jessica	9	27%	√	√	√	√	A
Kane	15	45%	D	√	B	√	√
Correena	25	76%	√	√	B	√	√
Arianna	24	73%	√	√	√	√	√

Item Bank Security Agreement

Educators using the Benchmark Interim Assessment Item Bank

may NOT;

- Add any items from other banks into the Benchmark Assessment Item Bank
- Modify the ETS items in any manner
- Export, copy, or save the items to any system, hardware, or software

Resources

Interim Assessment Home Page
http://oada.dadeschools.net/IAP/IAP.asp

Examview Test Generator
http://it.dadeschools.net/examview.htm

Edusoft
http://www.edusoft.com