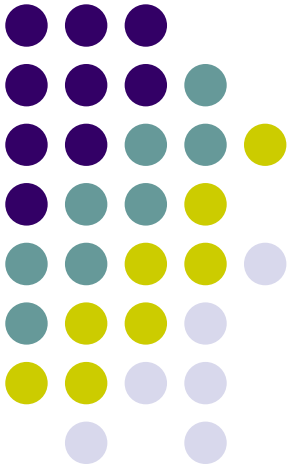
20th Century Achievement Constructs and Assessments: How Far to the 21st Century?

Panel discussion at the Center for Education Board Meeting, The National Academies

Steve Ferrara

American Institutes for Research

May 3, 2005



Overview

- Point of view
- Studies from an NSF-funded project
  - Method: Item Demands Analysis
  - Brief selected results
- Discussion
- Hope not to overstate (or understate) anything

Point of view

- Significant strides in last 25 years
  - Defining what achievement to assess
  - Designing, developing, and implementing achievement assessments
- Lost some ground in some content areas in the last 5 years

Point of view (cont.)

- State and other K-12 content area achievement tests aspire, but do not achieve complex achievement constructs (e.g., CC and SNP)
  - The achievement constructs they target (i.e., content standards)
  - Item demands
- You already know that

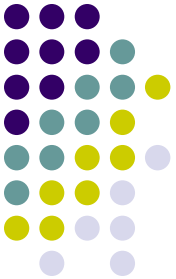
Study methods

- Item demands analysis (IDA)
  - Identify content area knowledge, content area skills, and broader processes that items elicit
  - Compare to cognitive targets
- IDA codes provide hypotheses
- Empirical support
  - Several other studies
  - Project cognitive labs for two state science assessments
  - Good rater agreement

Illustrative codes

- Use/apply
  - (e.g., Experimental design and scientific process skills; graphical representations given within the item or item set)
- Answer and explain scientific procedures, principles, or concepts
  - (e.g., Defend the answer given; explain why something happened or propose possible explanations)
- Analyze, categorize, hypothesize
- Create or invent
- Language-related demands
- Metacognition
- Visual-spatial analysis and use of graphics

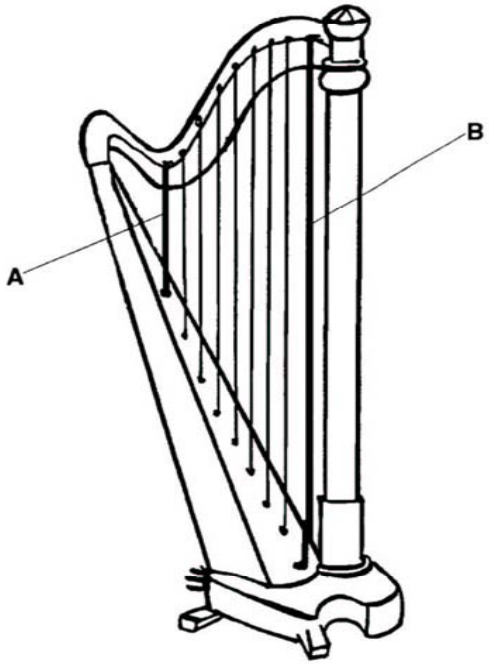
Grade 6 statewide science proficiency test (2002)



- Conventional design
  - 34 MC (55% of MPS), 10 SCR (32%), 2 ECR (13%)
- Content knowledge
  - Physical (30% of items), life (22%), earth/space science (15%), nature of science (33%)
- Levels of science processes
  - Acquiring (30% of items), processing (45%), and extending (25%) scientific knowledge

Misaligned: *Cynthia's Harp*

13. Cynthia plucks String A and String B in the harp shown below. String A makes a higher-pitched sound than String B makes. If Cynthia tightens String A and loosens String B, how will the sound each string makes compare with the sound it made before?



A. String A will still make a lower-pitched sound than String B.

B. String A and String B will both make a higher-pitched sound than they did before.

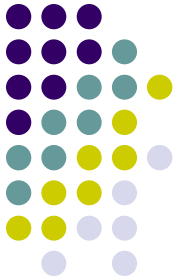
C. String A will make a higher-pitched sound than it did before, and String B will make a lower-pitched sound than it did before.

D. String A will make a lower-pitched sound than it did before, and String B will make a higher-pitched sound than it did before.

Misaligned: *Cynthia's Harp*

Cynthia's Harp (Physical Science, Processing Scientific Knowledge)			
	Evidence of Alignment with Target	**Evidence of Alignment with Other Science and Cognitive Targets**	**Evidence of Misalignment**
Content area knowledge	--	--	No science content knowledge observed (N=22)
Content area skills	Use/apply information given (N=1) Use/apply prior knowledge (N=21) Reiterate, re-explain, or summarize (N=1)	--	--
Broader cognitive processes	Use of graphics given (N=5)	--	No broader cognitive processes observed (N=9)

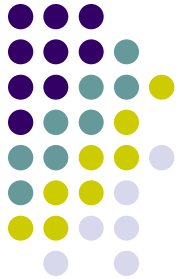
IDA main results (2005 paper)



- Two primary demands on examinees
  - Use and apply science declarative knowledge
  - Provide and explain answers

- Use/apply: 44 0f 46 total items
  - All 34 MC items
  - 8 of 10 SCR items
- Both ECRs
  - Use/apply and answer/explain

Science achievement construct achieved



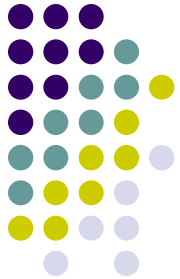
- Application of knowledge and understanding

- 25 of 46 items (54%)
  - Use information given with the item
- Mostly graphical information
  - Visual representations (*Shalonda's Balloon*; k=18)
  - Graphical representations (*Disasters Table*; k=5)
  - Textual information (*Marcia*; k=1)

Targeted and achieved constructs

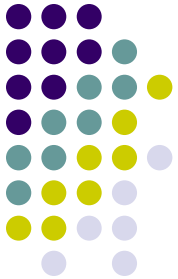
Targeted Level of Understanding	Expected and Observed Item Demand
Acquiring scientific knowledge	
Read graphs and tables; recall information	(11b, c, d) Use information given with the item (11h) Use prior knowledge and expectations
30% of items	(11b, c, d) 20 of 46 items (43%) (11h) 24 of 46 items (52%)
Processing scientific knowledge	
Organize and interpret information; recognize and infer relationships	(12a, b) Defend answer, explain why (13) Analyze, categorize, hypothesize
45% of items	(12a, b) 8 CR items (17%) (13) 0 items

Targeted and achieved constructs (cont.)



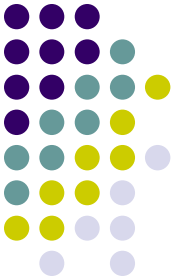
Targeted Level of Understanding	Expected and Observed Item Demand
Extending scientific knowledge	
Evaluate and transform ideas and hypotheses; propose solutions; draw conclusions; reason to solve problems	(11e) Use information generated with the item (13) Analyze, categorize, hypothesize (14) Create, invent (e.g., solutions)
25% of items	(11e) k=0 (13) k=4 (9%) (14) k=0

Achievement construct this assessment



- Requires more recall of science information than intended (52% vs. 30%)
  - Additional 20 items (43%): use given information
- Requires less organizing and interpreting information and recognizing and inferring relationships than intended (17% vs. 45%)
- Requires less evaluating, problem solving, and reasoning than intended (9% vs. 25%)

Maryland School Performance Assessment Program (MSPAP)



- 1991-2001
- All constructed-response, coherent assessment tasks
- Hands-on science investigations
- Satisfactory level by 2000
  - 26-54% of students across 24 school systems, three grades, six content areas
    - 15-38% of Black students
    - 23-45% of Hispanic students
  - 83 schools Satisfactory composite, up from 11 in 1993

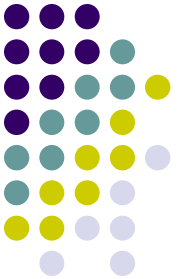
Bouncing Light

- Review an experiment with hinged mirrors and number of reflections
  - Form a hypothesis
- Conduct the experiment for various angles between the mirrors
  - Graph results from a group investigation
  - Evaluate the earlier hypothesis
- Using the results
  - Design a hall of mirrors
  - Make a penny “disappear”

IDA results: MSPAP task

- History and Nature of Science
  - (not Physical Science)
- Science skills
  - Use graphical representations given with the item
  - Analyze, categorize, hypothesize

Definitions of complex achievement constructs



- Illuminate distinctions and relationships among concepts, skills, and processes
- Describe similarities and differences from other achievement constructs
- Specify, realistically, thinking processes to be assessed
- Reflect the development of proficiency

- (From *KWSK*, NRC, 2001, p. 45, “targets of inference”)

Concluding comments (1)

- State content standards do not define complex achievement constructs
- Typical/many state assessments do not target complex achievement constructs

- A matter of conceptual distance

Concluding comments (2)

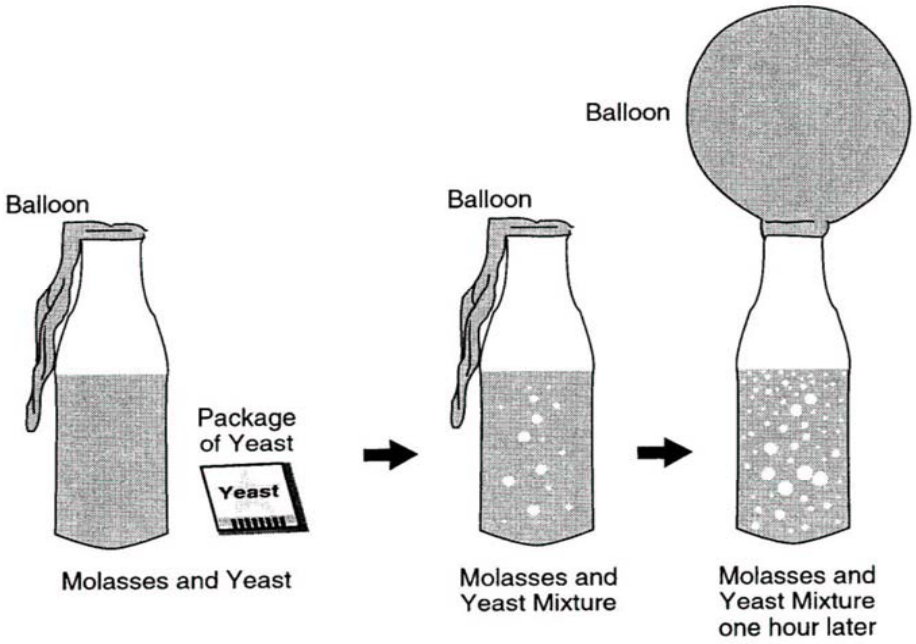
- NCLB a resource hog
  - (Even conceding good intentions)
  - Costs of more complex assessments
- Political will in the face of slow growth

- A matter of resources and perseverance

- Intentionally blank

Partially aligned: *Molasses and Yeast*

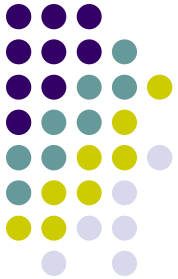
Directions: Use the pictures below to answer question 35.



35. The pictures above show what happens when molasses and yeast are mixed and allowed to sit. When mixed together, the yeast eats the molasses and produces carbon dioxide and alcohol. In part 1 of your **Answer Booklet** describe what physical change you see.

In part 2 explain what chemical change has occurred.

Partially aligned: *Molasses and Yeast*



Molasses & Yeast (Physical Science, Extending Scientific Knowledge)			
	Evidence of Alignment with Target	**Evidence of Alignment with Other Science and Cognitive Targets**	**Evidence of Misalignment**
Content area knowledge	Physical science (N=16)	Life science (N=2)	No science content knowledge observed (N=4) Lack of necessary content area knowledge (N=1)
Content area skills	Use/apply information given (N=9) Use/apply prior knowledge (N=1) Defend the answer given (N=1) Reiterate, re-explain, or summarize (N=1) Explain why something happened (N=4) Provide an explanation (N=5)	Describe patterns in data, procedures, or results (N=11)	--
Broader cognitive processes	Use of graphics given (N=21)	--	No broader cognitive processes observed (N=1)