Trends in Science Assessment: Linking Assessment Methods to Facets of Achievement

Richard J. Shavelson
Stanford University

Invited Talk
Department of Educational Measurement
Umea University
May 17, 2000
Overview of Talk

• What does it mean to “achieve” in science?
• How might we measure achievement in science?
• What evidence is there that we’re measuring different aspects of achievement?
• What remains (or what doesn’t remain!) to be done?
Toward An Achievement Framework: Knowledge Components

Characteristics That Vary According to Proficiency Level

Extent
(How much?)

Structure
(How is it organized?)

Others
(Precision? Efficiency? Automaticity?)

Declarative Knowledge
(Knowing the “that”)

- Domain-specific
  content:
  - facts
  - concepts
  - principles

Procedural Knowledge
(Knowing the “how”)

- Production
  system--
  condition-
  action rules

Strategic Knowledge
(Knowing the “which,” “when,” and “why”)

- Problem schemata/
  strategies/
  operation systems

Cognitive Tools:
Planning
Monitoring
Linking Assessments to Achievement Components

<table>
<thead>
<tr>
<th>Extent</th>
<th>Declarative Knowledge</th>
<th>Procedural Knowledge</th>
<th>Strategic Knowledge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structure</td>
<td>Multiple-Choice</td>
<td>Performance Assessments</td>
<td>Performance Assessments</td>
</tr>
<tr>
<td></td>
<td>Conceptual Maps</td>
<td>Procedural Maps</td>
<td>Models/Mental Maps</td>
</tr>
<tr>
<td>Others</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Air is made up of many gases. Which gas is found in the greatest amount?

A. Nitrogen
B. Oxygen
C. Carbon Dioxide
D. Hydrogen
Structure of Declarative Knowledge: A Concept Map of a Concept Map!

- **Concept Map**
  - is a graph
    - consist of nodes labeled concepts
      - represent
    - linked by labeled lines
      - denote a relation
    - consist of concepts between a pair of
Structure of Declarative Knowledge: Eleven-Year-Old’s Concept Map

From White & Gunstone: “Probing Understanding” (1992, p. 16)
Procedural Knowledge: Performance Assessment: TIMSS Pop. 2

**PULSE**

**At this station you should have**
- A watch
- A step on the floor to climb on

**Read ALL directions carefully.**

**Your task:**
Find out how your pulse changes when you climb up and down on a step for 5 minutes.

**This is what you should do:**
- Find your pulse and be sure you know how to count it. IF YOU CANNOT FIND YOUR PULSE ASK A TEACHER FOR HELP
- Decide how often you will take measurements starting from when you are rest.
- Climb the step for about 5 minutes and measure your pulse at regular intervals.

1. Make a table and write down the times at which you measured your pulse and the measurements you made.
2. How did your pulse change during the exercise?
3. Why do you think your pulse changed in this way?
Procedural Knowledge: Performance Assessments for FOSS Curriculum
### Types of Tasks and Scoring Systems

<table>
<thead>
<tr>
<th>Type of Assessment Task</th>
<th>Scoring System</th>
<th>Component Identification</th>
<th>Classification</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comparative Investigation</td>
<td>Analytic</td>
<td>Paper Towels</td>
<td>• Bugs</td>
<td>Others</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Incline Planes</td>
<td>• Friction</td>
<td>Others</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Electric Mysteries</td>
<td>• Bubbles</td>
<td>Others</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rocks and Charts</td>
<td>• Mystery Powders</td>
<td>Others</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Day-Time Astronomy</td>
<td></td>
<td>Others</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Others</td>
</tr>
<tr>
<td>Dimension-Based</td>
<td>Holistic Rubric</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Strategic Knowledge: Mental Models(?)

• What causes day and night?
  A. **The earth spins on its axis** (.66)
  B. **The earth moves around the sun** (.26)
  C. **Clouds block out the sun** (.03)
  D. **the sun goes round the earth** (.04)

Strategic Knowledge: Mental Models

• (A) A rocket is moving along sideways in deep space, with its engine off, from point A to point B. It is not near any planets or other outside forces. Its engine is fired at point B and left on for 2 sec while the rocket travels from point B to point C. Draw in the shape of the path from B to C. (Show your best guess for this problem even if you are unsure of the answer.)

• (B) Show the path from C after the engine is turned off on the same drawing.

Some Empirical Evidence: Links between Knowledge and Measurement Methods

Correlations from Shultz’s Dissertation (N=109 6th Graders Studying Ecology):
– CTBS Reading and M-C: 0.69
– Reading and Concept Map: 0.53
– M-C and CM: 0.60
– Reading and Performance Assessment: 0.25
– M-C and PA: 0.33
– CM and PA: 0.43
Some Empirical Evidence: Other Evidence

• Correlations (covariances) tell you about relative rank ordering on different measures--that’s only part of the story.

• Evidence is needed for the cognitive claims that different methods measure somewhat different aspects of achievement:
  – Talk aloud
  – Focus group
  – Group work
  – Other?
Cognitive Validity Framework

Comparing assessment tasks on three dimensions

Intended Task Demands -> Inferred Cognitive Activity -> Scores Obtained

Congruence
Construct a concept map that reflects what you know about what a concept map is. Organize the terms in relation to one another in any way you want.
Examine the map, the blank nodes and the terms provided on the list below. Select the term that corresponds to each node and write it down inside the circle. Use each term only once.

**List of Concepts:**
- concepts
- graph
- labeled lines
- propositions

**Concept Map**

- *concept map* is a
  - *nodes*
    - linked by *represent*
    - *linked by*
      - *that describe the relationship between two nodes are called*
Talk Aloud with Concept Maps: Method Variation

Correspondence Between Directedness and Inferred Activity

- **Proportion Scores**
- **Verbal Codes**
  - Explanation
  - Monitoring
  - Conceptual Errors
  - No-Code

**Graph Legend**
- Green: Construct-A-Map
- Blue: Fill-In-Nodes
- Light Blue: Fill-In-Lines

- Low Directed
- High Directed
Future Directions

• Does the distinction between knowledge types and a working definition of achievement make any sense?
• How defensible are the links between measurement methods and knowledge types?
• What other measurement methods tap into these knowledge types?
• How can we measure mental models on a large scale?