Recent Developments Linking (Science) Teaching, Learning & Assessment

Richard J. Shavelson
Stanford University

Invited Address
European Association for Research on Learning and Instruction
Fribourg, Switzerland
August 29, 2001
The Real Authors

- Carlos Ayala
- Gail Baxter
- Min Li
- Anders Rosenquist
- Maria Araceli Ruiz-Primo
- Guillermo Solano-Flores
- Susan Schultz
- Yin Yue
Overview

• Advances made in
  – Cognitive theory
  – Assessment methods
  – Interpretation

• Advances made at both summative and formative assessment levels

• Gap between teaching and assessment (in the U.S.) must be closed to improve learning
The Assessment Triangle

Learning/Achievement
(cognition)

Observation
(assessment)

Interpretation

# Advances in Cognition

## I. Summative-level

<table>
<thead>
<tr>
<th>Type of Knowledge</th>
<th>Definition</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Declarative</td>
<td>Knowing that</td>
<td>Concepts &amp; facts</td>
</tr>
<tr>
<td></td>
<td></td>
<td>“What is saturation?”</td>
</tr>
<tr>
<td>Procedural</td>
<td>Knowing how</td>
<td>Actions, steps, &amp; procedures</td>
</tr>
<tr>
<td></td>
<td></td>
<td>“Saturate 50ml water with salt”</td>
</tr>
<tr>
<td>Schematic</td>
<td>Knowing why</td>
<td>Principles &amp; mental model solution</td>
</tr>
<tr>
<td></td>
<td></td>
<td>“Can saturated salt solution dissolve more sugar? Why or why not?”</td>
</tr>
<tr>
<td>Strategic</td>
<td>Knowing about</td>
<td>Strategies &amp; cognitive operations</td>
</tr>
<tr>
<td></td>
<td>knowing</td>
<td>---</td>
</tr>
</tbody>
</table>
Advances in Assessment
I. Summative-level

- Multiple-choice/Item response theory
- Constructed response/”E-rater”
- Concept/cognitive mapping/(Bayesian) networks
- Performance assessment/G theory
Assessing Declarative Knowledge Extent
Multiple-Choice: TIMSS Pop. 2

Air is made up of many gases. Which gas is found in the greatest amount?

A. Nitrogen
B. Oxygen
C. Carbon Dioxide
D. Hydrogen
Assessing Declarative Knowledge Structure

Concept Map: Eleven-Year-Old’s Map

Assessing Procedural Knowledge
Performance Assessment: *Daytime Astronomy*

Students are asked to model the path of the sun from sunrise to sunset and use direction, length, and angles of shadows to solve location problems.

*Figure Removed To Reduce Size Of File*
Assessing 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?
Assessing Strategic Knowledge Selection: *Mental Models in Physics*

- (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.

Assessing Strategic Knowledge
Multiple-Choice: *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)

Advances in Interpretation

I. Summative-level

<table>
<thead>
<tr>
<th>Assessment of Knowledge</th>
<th>Declarative Knowledge</th>
<th>Procedural Knowledge</th>
<th>Strategic Knowledge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extent</td>
<td>• Multiple-Choice</td>
<td>• Performance Assessments</td>
<td>• Multiple-Choice</td>
</tr>
<tr>
<td></td>
<td>• Fill-in</td>
<td>• Journals</td>
<td>• Interviews</td>
</tr>
<tr>
<td></td>
<td>• Journals</td>
<td></td>
<td>• Performance Assessment</td>
</tr>
<tr>
<td>Structure</td>
<td>• Concept Maps</td>
<td>• Procedure Maps</td>
<td>• Multiple-Choice</td>
</tr>
<tr>
<td></td>
<td>• Cognitive Maps</td>
<td></td>
<td>• Short Answer</td>
</tr>
<tr>
<td>Others</td>
<td></td>
<td></td>
<td>• Performance Assessment</td>
</tr>
</tbody>
</table>
## Links between Knowledge Types and Assessment Methods: TIMSS-R

<table>
<thead>
<tr>
<th>Type of Knowledge</th>
<th>Multiple-Choice</th>
<th>Open-Ended</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analytic (Declarative &amp; Schematic)</td>
<td>10</td>
<td>12</td>
<td>22</td>
</tr>
<tr>
<td>Procedural</td>
<td>6</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Conceptual (Declarative)</td>
<td>16</td>
<td>9</td>
<td>25</td>
</tr>
<tr>
<td>Factual (Declarative)</td>
<td>38</td>
<td>1</td>
<td>39</td>
</tr>
<tr>
<td>“Not Science”</td>
<td>4</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>73</strong></td>
<td><strong>27</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>

Empirical Evidence of Link: Statistical & Cognitive

- Statistical fit: $\chi^2=357.47$, df=333, $p=.17$, CFI=.999
- Cognitive—talk alouds

Advances in Cognition
II. Formative-Level

• Importance of prior knowledge, practice and feedback, transfer, social context

• Expert-novice differences in specific knowledge domains
  – Approach (principled/trial & error)
  – Schema (rich/weak-disconnected)
  – Mental-models develop from naïve scientifically justifiable (through multiple paths)

• Role of meta-cognition
  – Knowing one’s cognitive capabilities
  – Stepping back from problem-solving activity to evaluate
Mental Models of the Earth

“For each identified earth shape model we generated the pattern of responses expected if the child had used this model consistently to answer our questions…” (p. 550).

Advances in Assessment

II. Formative-Level

• Students’ learning activities (individual & group work)
• Students’ projects
• Students’ science notebooks
• Teachers’ observations
• Teachers’ assessments
II. Formative-Level

- Classical test theory
- Generalizability theory
- Item response theory
- Multi-attribute (multidimensional IRT) models
- Latent class models
- Bayesian net models
Advances in Assessment of Learning—But the Teaching Gap

- Teachers hold strong beliefs about assessment… detached from the teaching act
- Teachers’ assessment practices fit this belief—assessment used for grading and defending grades to parents
- Teachers rarely take time to give qualitative feedback to students to close the learning gap
- Indeed, as often as not, teacher feedback is inaccurate
- *The best assessments cannot help foster student learning without changes in teachers’ “mindframes”*
Formative & Summative Assessment:
Degree of Instructional Sensitivity

**Depth of Assessment Probe**

**Remote:** Standardized National Science Achievement Tests

**Distal:** Large-Scale Performance Assessment from State/National Curriculum Framework

**Proximal:** Same Concept/Principle--New Investigation

**Close:** “Embedded” Assessments -- A Slightly More Advanced Activity in Unit

**Immediate:** Lab Notebooks & Classroom Tests

Classroom Instruction
Formative & Summative Assessment

<table>
<thead>
<tr>
<th>Type</th>
<th>Purpose</th>
<th>Agency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formative</td>
<td>Learning</td>
<td>student, teacher</td>
</tr>
<tr>
<td></td>
<td>Certification</td>
<td>teacher, individual</td>
</tr>
<tr>
<td></td>
<td>Accountability</td>
<td>external tests, sample surveys</td>
</tr>
<tr>
<td>Summative</td>
<td>Match or Mismatch</td>
<td>external tests, individual</td>
</tr>
</tbody>
</table>

Paul Black 3/98
Formative Assessment Methods

• Teacher
  – Observations and group talk
  – Questions
  – Interviews
  – Projects
  – Journals
  – Curriculum-provided and/or teacher assessments
• Self and Peer (*Clear goals absolutely essential*)
  – Review/grade each other’s work
  – Review each other’s journals
  – Reflect on learning
• Intelligent tutoring
Formative Assessment Methods

• Teacher
  – Observations and group talk
  – Questions
  – Interviews
  – Projects
  – Journals
  – Curriculum-provided and/or teacher assessments

• Self and Peer (*Clear goals absolutely essential*)
  – Review/grade each other’s work
  – Review each other’s journals
  – Reflect on learning

• Intelligent tutoring
Science Notebooks: Opportunity for Formative Assessment

Figure Removed To
Reduce Size Of File
Some Findings About Students’ Journals as Assessment Tools

Reliability:
• Raters can consistently identify journal entries
• Students’ science journals can be reliably scored

Validity:
• Inferences about implementation using journal scores were justified
• Inferences about students’ performance were also justified

Usefulness:
• Unit implementation and teacher feedback scores helped to explain differences in the performance across classrooms
Concluding Comments

- Substantial advances have been made in cognition, assessment methods, and a chain of logical reasoning from data to cognitive model with new psychometrics.
- The best cognitive and psychometric theory is no match for teachers’ beliefs about the role of assessment in teaching.
- We need to tighten the link between cognition, assessment and teaching to improve student learning.