On Linking Learning, Assessment, and Interpretation

Min Li

Stanford University

Invited Talk
University of Washington
April 26th, 2001
Overview of Talk

- Conceptual framework: The assessment triangle
- Dissertation study – evaluating the links between science achievement to assessments
- Classroom assessment study – looking into students’ learning in the classroom through Science journals
- Directions and implications
Conceptual Framework:
The Assessment Triangle
The Assessment Triangle

Learning/Achievement
(cognition)

Assessment

Interpretation

Pellegrino, Chudowsky, & Glaser, in press
Dissertation Study: Evaluating the Link between Science Achievement and Assessments
The Assessment Triangle: Dissertation Study

Science achievement as four types of knowledge:
- Declarative
- Procedural
- Schematic
- Strategic

Assessment methods:
- Multiple-choice
- Free-response
- Performance-assessment

Evaluate Interpretations:
- Logically
- Empirically
Defining Science Achievement

Science Achievement

- Declarative
- Procedural
- Schematic
- Strategic

- Definitions, terms & facts
- Steps, actions, & algorithms
- Theories & mental models
- Strategies & conditional knowledge
# Linking Knowledge Types to Assessment Methods

<table>
<thead>
<tr>
<th>Knowledge Type</th>
<th>Assessment Method</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Declarative</strong></td>
<td>Multiple-choice, free-response</td>
</tr>
<tr>
<td><em>Knowing that</em></td>
<td></td>
</tr>
<tr>
<td><strong>Procedural</strong></td>
<td>Performance assessment</td>
</tr>
<tr>
<td><em>Knowing how</em></td>
<td></td>
</tr>
<tr>
<td><strong>Schematic</strong></td>
<td>Multiple-choice, free-response, concept-mapping</td>
</tr>
<tr>
<td><em>Knowing why</em></td>
<td></td>
</tr>
<tr>
<td><strong>Strategic</strong></td>
<td>Performance assessment with an open structure</td>
</tr>
<tr>
<td><em>Knowing about knowing</em></td>
<td></td>
</tr>
</tbody>
</table>
Method: Sample

• TIMSS science items (American Pop 2)
  – Booklet 8
    • 22 multiple-choice
    • 10 free-response
  – Performance assessment tasks

• Ten experts selected for think-aloud study
  – 5 physics and 5 biology graduate students
Method: Procedure

The link was evaluated logically and empirically:

- Logical analysis
  - Coding the characteristics of items and linking items to knowledge types

- Empirical analysis
  - Exploratory and confirmatory factor analysis to examine the underlying covariance patterns
  - Think-aloud study to infer students’ cognition
## Logical Analysis:

**Classification of TIMSS Items by Knowledge Type**

<table>
<thead>
<tr>
<th>Knowledge Type</th>
<th>Number of Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Declarative</td>
<td>13</td>
</tr>
<tr>
<td>Schematic</td>
<td>6</td>
</tr>
<tr>
<td>Procedural</td>
<td>5</td>
</tr>
<tr>
<td>Strategic</td>
<td>0</td>
</tr>
<tr>
<td>?</td>
<td>4</td>
</tr>
</tbody>
</table>

- **Declarative** (3 items): 13
- **Schematic** (1 item): 6
- **Procedural** (1 item): 5
### P4. What happens when an animal hibernates?

A. There is no life in any of its parts.  
B. It stops breathing.  
C. Its temperature is higher than when it is active.  
D. It is absorbing energy for use when it is active.  
E. It is using less energy than when it is active.
Example: A Procedural-knowledge Item

P1. The graph shows the progress made by an ant moving along a straight line.

If the ant keeps moving at the same speed, how far will it have traveled at the end of 30 seconds?

A. 5cm
B. 6cm
C. 20cm
D. 30cm
Example: A Schematic-knowledge Item

P5. The water in a tube is heated, as shown in the diagram. As the water is heated, the balloon increases in size. Explain why.
Factor Analysis

A good statistical fit indicated by fit measures:
- \( \chi^2 = 357.47 \),
  \( \text{df} = 333 \), \( P = .17 \)
- CFI = .999
Factor Analysis

- Knowledge-type items clustered together, showing significant loadings on the three knowledge factors.
- Declarative, procedural, and schematic knowledge factors were highly correlated.
- Comparison with alternative models (e.g., one general factor) favored the knowledge-factor model.
Protocol Analysis

Examples of Experts’ Verbalizations

The declarative-knowledge item

P4. What happens when an animal hibernates?

A. There is no life in any of its parts.
B. It stops breathing.
C. Its temperature is higher than when it is active.
D. It is absorbing energy for use when it is active.
E. It is using less energy than when it is active.
Protocol Analysis

Examples of Experts’ Verbalizations

**Expert 1:** “What happens when an animal hibernates? Okay, what I know, hibernation means, sleeps for a long time and heart rate slows down …”

**Expert 2:** “The answer is just said there (from reading the statement) … all the process slows down. Well, I know when animals hibernate, they lay down and they do not use too much energy.”
P5. The water in a tube is heated, as shown in the diagram. As the water is heated, the balloon increases in size. Explain why.
Protocol Analysis

Examples of Experts’ Verbalizations

Expert 1: “… when it is hotter, the atoms move faster. So they evaporate, so there is more vapor. That vapor, that is more pressure on the balloon…”

Expert 2: “The balloon increases because of pressure, that is what causes the balloon expands. And also, when you heat something, even without water, when you heat gas, pressure of the volume tends to increase.”
Conclusions of Dissertation Study

- Logical and factor analyses supported the distinctions between knowledge types.
- Protocol analysis revealed differences in use of knowledge types partially due to item characteristics.
Classroom Assessment Study: Looking into Students’ Learning in the Classroom through Science Journals
Science journal is a compilation of entries that provides a partial record of the instructional experiences a student had in her classroom during a certain period of time.

**SEPARATING MIXTURES**

<table>
<thead>
<tr>
<th>Gravel</th>
<th>Color</th>
<th>Texture</th>
<th>Particle shape</th>
<th>Particle size</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>dark grey, round, hard, metal</td>
<td>rough</td>
<td>round</td>
<td>3 mm</td>
<td>different sizes</td>
</tr>
<tr>
<td>Powder (Diatomaceous earth)</td>
<td>Snow white, soft</td>
<td>powdery, round</td>
<td>how can creat a cone</td>
<td>estimation of 3 mm</td>
<td>12</td>
</tr>
<tr>
<td>Salt (Sodium chloride)</td>
<td>grey, cloudy</td>
<td>gritty, messy</td>
<td>hexagonal</td>
<td>looks like fine sugar</td>
<td>12</td>
</tr>
</tbody>
</table>

**Tested experiment 2**

**Testing weight**

We are going to change the weight by adding a penny.

I think weight will change the number of cycles because the weight will be pulling the string down with it.

Practiced estimation: 12

Actual count: 12

Absolutely right! Weight is not a variable because it doesn't change the outcome.
The Assessment Triangle: Science Journals

Learning/Achievement

- Student performance
  - Scientific communication
  - Conceptual understanding
  - Procedural understanding

- Opportunities to learn
  - Instructional implementation
  - Quality of teacher feedback

Journals as Assessment Tools:

- at the individual level and at the aggregated classroom level.
- an immediate/unobtrusive assessment
The Assessment Triangle: Science Journals

- Can science journals provide trustworthy and valid evidence on student performance?
- What do journals tell us about student performance?
- What do journals tell us about opportunity to learn?
Method

• Sample
  – 10 fifth grade classrooms
  – Two Full Option Science System (FOSS) units, Variables in the fall and Mixtures in the spring, were taught.
  – A random stratified sample from each class: 2 low, 2 middle, and 2 high
Method

• Coding
  – Each entry was coded into different scores:
    • Instructional implementation
    • Type of entry
    • Student performance
    • Teacher feedback
  – An analytic coding criteria defined the values.
Method

• Procedures
  – Pre-posttest design using performance assessments
  – 28 Variables and 22 Mixtures journals were coded by two coders.
## Technical Characteristics of Journal Scores

### Reliability

<table>
<thead>
<tr>
<th>Type of entry</th>
<th>% of Agreement</th>
<th>Intercoder Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Student performance</td>
</tr>
<tr>
<td>Variables</td>
<td>85</td>
<td>.85</td>
</tr>
<tr>
<td>Mixtures</td>
<td>85</td>
<td>.84</td>
</tr>
</tbody>
</table>

### Validity

Students’ journal scores were correlated with their performance assessment scores (on average $r = .52$).
Student Performance

- **Variables in the fall**: Blue bars
- **Mixtures in the spring**: Pink bars

<table>
<thead>
<tr>
<th>Category</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scientific communication</td>
<td>1.5</td>
</tr>
<tr>
<td>Conceptual understanding</td>
<td>2.0</td>
</tr>
<tr>
<td>Procedural understanding</td>
<td>1.5</td>
</tr>
</tbody>
</table>
Opportunity to Learn: Learning Activities

- Defining
- Exemplifying
- Applying concepts

Variables
Mixtures

Percentage

0 5 10 15 20 25
Opportunity to Learn: Learning Activities

- Defining
- Exemplifying
- Applying

Concepts

Variables

Mixtures

Percentage

Predicting

Results

Interpreting

Res & interpret

Procedures

Experiments

Designing
Opportunity to Learn: Teacher Feedback

- Teachers did not provide feedback despite errors or misconceptions that were evident in the students’ journals.
- Only 4 among the 10 teachers provided feedback!
Conclusions of Classroom Assessment Study

• Science journals can be reliably scored and be used as a valid assessment tool.

• Students did poorly in scientific communication and showed partial science understanding in their journals.

• Most teachers did not effectively use science journals.

• Teachers had very limited content knowledge. They did not know how to promote or assess student learning.
Directions and Implications
Directions and Implications

Dissertation Study

- What other assessment methods can tap into these knowledge types?
- How can we measure types of knowledge at the classroom level?
- How can we use the distinctions between knowledge types to improve teachers’ practice?
- How can we use these distinctions to promote students’ learning?
Directions and Implications

*Classroom Assessment Study*

- Can journals be used to gather information on teaching for the accountability purposes?
- How can we make the use of journals relevant to and practical for teachers in their practice?
- How can we use these findings to improve teachers’ classroom assessment practices?
Directions and Implications

*My Research on Assessment*

Learning/Achievement
(cognition)

Assessment

Interpretation
THANK YOU!

I look forward to your questions and comments.