How good is our evidence of student knowledge: Some cognitive issues in diagnostic assessment

Slava Kalyuga

Educational Assessment Australia
University of New South Wales

Roundtable 2004
Knowing what students know: The science and design of educational assessment.


From Executive Summary (pp. 3-4):
A **model of cognition** and learning should serve as the cornerstone of the assessment design process…

Understanding the contents of **long-term memory** is especially critical for determining what people know…

Assessment should evaluate what **schemas** an individual has and under what circumstances he or she regards the information as relevant…

Studies of **expert-novice differences** in subject domains illuminate critical features of proficiency that should be the targets for assessment.
Outline:

Review of our cognitive architecture: Why do we need diagnostic assessment of organized knowledge structures?

Are traditional tests sufficient for diagnostic assessment?

Some suggested non-traditional approaches
The need for cognitively-based approaches:

- Mislevy (1994; 1996)
- Pellegrino, Baxter, & Glaser (1999)
- Snow & Lohman (1989)

Linking cognitive science and psychometrics:

- Tatsuoka (1990)

Assessment of cognitive constructs:

- Masters & Forster (1996)
Suppose 5 days after the day before yesterday is Friday. What day of the week is tomorrow?

Dial a number from a telephone directory.

How many windows are in your house/unit?
Working Memory

Constructing mental representations of a situation or task

Long-Term Memory

Knowledge base

Sensory Memory:

Incoming information
Cognitive studies of expertise:

knowledge base in LTM is central to cognitive processing
Chess expertise


Grand masters are better than weekend players in reproducing briefly presented chess positions from real games.

Knowledge of large numbers of different game configurations held in LTM dramatically altered the characteristics of working memory
Arithmetic word problem schemas

(Marshall, 1995)

Change, Group, Compare, Restate, and Vary schemas
After 6 passengers had left the bus, 9 passengers remained. How many passengers were on the bus initially? (Change)

Peter's book contains 50 pages. Peter read 15 pages in the morning. In the afternoon, he read the remaining pages and finished the book. How many pages did Peter read in the afternoon? (Group)
Reading comprehension
(linguistic schemas)

This was the course that the student, whom the school that was criticized by the newspaper expelled, failed.
This was the course that the student whom the school that was criticized by the newspaper expelled failed.
Knowledge base in Long-Term Memory affects the way we process information in Working Memory:

WM is very limited when dealing with novel information

WM has no known limits when dealing with information that has been organized and stored in LTM
Organized knowledge structures in LTM affect the way we solve problems

Novices:
weak problem-solving methods

Experts:
rapid retrieval and application of previously acquired organized knowledge structures
Solve for x: \( 5x = -4 \)
Solve for $x$: $5x = -4$

$x = -\frac{4}{5}$
Solve for $x$: $5x = -4$

$rac{5x}{5} = -\frac{4}{5}$
Solve for x: \[ 5x = -4 \]

\[
\begin{align*}
5x &\quad -1 = -5 \\
5x &\quad -1.5 = -7.5 \\
5x &\quad -2.25 = -12.25 \\
5x &\quad -5 = -2.5 \\
5x &\quad -12.5 = \\
5x &\quad -1.25 = -6.25 \\
5x &\quad 5.5 = -2.75 \\
5x &\quad 60 = -3 \\
5x &\quad .7 = -1.5 \\
5x &\quad 75 = 3.25 \\
5x &\quad 8 = -5 \\
\end{align*}
\]

\[ x = -0.8 \]

Solve for x: \[ -4x = -6 \]

\[
\begin{align*}
-4x &\quad 2 = -8 \\
-4x &\quad 1.3 = -5.2 \\
-4x &\quad 2 = -8 \\
-4x &\quad 1.5 = -6 \\
\end{align*}
\]

\[ x = 1.5 \]
Diagnostic cognitive assessment:

- evidence based on remote results of cognitive processes
  vs

- evidence based on immediate traces of cognitive processes
Diagnostic cognitive assessment:

typical time scale of cognitive processes (retrieval of learned knowledge from LTM, carrying out simple transformations, etc.): up to several seconds.
Diagnostic cognitive assessment

- sensitive to different cognitive attributes

- sensitive to different levels of proficiency

- practically usable
Expertise reversal effect

Instructional designs or procedures that are effective for novices may be ineffective for more proficient (expert) learners.

Overview in:
Assume you wish to determine the appropriate R.P.M. to drill a 25 mm diameter hole in the bronze workpiece.

Step 1. Select the cutting speed.

Step 2. Select the diagonal line.

Step 3. Select the vertical line.

Step 4. Find the intersection point.

Step 5. Select the horizontal line.

Step 6. Read off the R.P.M.
Instructional implications:

Instructional techniques need to change radically with alterations in learner expertise.

(Paradox: less is more!)

It is critical to have a simple rapid measure of learner proficiency in a domain.
The theory:

If knowledge structures in LTM alter the characteristics of WM, then tests of WM content provide a measure of levels of expertise.

Such tests will effectively assess the extent to which WM limits have been altered by knowledge structures in LTM.

Long-Term Working Memory: sufficiently durable and interference proof.
Diagnostic approach (general):

What is the highest level of organised knowledge structures (if any) a person is capable of retrieving and applying to the briefly presented material?
Diagnostic approach:

Presenting learners with a task for a limited time and asking them to indicate their *first step* towards solution.

\[
\frac{2x + 1}{3} = 1
\]

A: multiply both sides by 3

\[
2x + 1 = 3
\]

B: subtract 1 from both sides

\[
2x = 2
\]

C: divide both sides by 2

\[
x = 1
\]
\[
\begin{align*}
2x + 1 &= 3 \\
3x + 1 &= 2 \\
4x &= 1 \\
3x &= 1 \\
2x &= 2 \\
\end{align*}
\]
N-step task:

traditional test: $\frac{N(N+1)}{2}$ steps

rapid test: $N$ steps

Reduction of testing time by a factor of $\frac{(N+1)}{2}$ (with a simplifying assumption of equal times for each step)
Skipping steps reflects a higher level of proficiency: the learner has corresponding operations automated or is able to perform these operations mentally without writing them down.
Scoring method

If a learner omitted some intermediate steps, she/he was allocated an additional score for each skipped step.

For $2x + 1 = 3$ scores 0 to 4:

4: $x = 1$

3: $2x/2 = 2/2$

2: $2x = 2$

1: $2x + 1 - 1 = 3 - 1$
Validation studies

Algebra Y 9-10, N=45:
- test time reduced by a factor of 4.9;
- correlation with a traditional test .92.

Coordinate geometry Y 9, N=20:
- test time reduced by a factor of 2.5;
- correlation with a traditional test .85

Kalyuga & Sweller (2004)
First-step diagnostic method could be more rapid and more sensitive to underlying knowledge structures than traditional approaches.

(Paradox: less is more!)
General procedure:

1. **Cognitive model** (knowledge structures to be assessed)
2. **Pattern of tasks** (to provide evidence about these knowledge structures)
3. **Evidence model** (scoring procedure, measurement model)
Arithmetic word problems

Change, Group, Compare, Restate, and Vary schemas
There were 20 passengers on the lower deck of a boat. The number of people on the upper deck was 14 more than the number of people on the lower deck. How many passengers were on the boat? (Change + Group)
On the following pages, you will see 20 arithmetic problems. For each problem, rapidly write only the first step that you would write if you were solving this problem.

For example, when asked to solve the problem Mary had 3 books. She bought 2 more books. Then she bought another book. How many books does Mary have now? some people would first write 3 + 2 or 5, others could start from 3 + 2 + 1 or 5 + 1, and some might immediately write the final answer 6.
Validation study

Y 8, N= 55 :

test time reduced by a factor of 2.8;
correlation with a traditional test .72


Multidimensional Rash model (MRCML: Adams, Wilson, & Wang, 1997), ConQuest software, partial credit model (Wu, Adams, & Wilson, 1998);
Bayesian conditional probabilities estimations (MCMC procedure), WinBUGS software (Spiegelhalter et al, 2003)
Rapid verification technique

Presenting learners with a series of possible task solutions for a limited time and asking them to rapidly verify the suggested solution steps.
The artist, who performed for the crowd that gathered to enjoy the show, left.

1 local thematic violation

3 sentence nodes to parse at one time
Working memory capacity allows no more than 4 local violations (Gibson, 1998) or 2 sentence nodes or clauses to parse at one time (Kimball, 1973).

Cognitive difficulty

\[ D = \max \{ \text{local thematic violations; sentence nodes or clauses to parse simultaneously} \} \]
The artist, who performed for the crowd that gathered to enjoy the show, left.

- The artist left.

- The artist enjoyed the show.

- The crowd gathered for the show.

- The crowd left.
This was the course that the student, whom the school that was criticized by the newspaper expelled, failed.

\[ D = \max \{7; 4\} \]
The school was criticized by the student.

The school expelled the student.

The course was criticized by the newspaper.

The newspaper criticized the school.
Validation study

rapid test: 18 sentences/ 72 responses

traditional test: 8 passages, 42 multiple choice questions

Y 7, N= 19:

test time reduced by a factor of 3.7

correlation with a traditional test  .63

reliability (Cronbach’s alpha)   .71
A ship is traveling at 7 m/s. A dog runs across the deck at the same speed in a direction of 60° relative to the direction of motion of the ship. What is the velocity of the dog relative to the sea?
7 m/s

Right

Wrong

Don't know
Validation study

rapid test:
25 tasks presented for 20 s each;
5 categories;
from 6 to 2 verification statements

Y 11, N= 23:
test time 16.5 min
reliability (Cronbach’s alpha)  .76
The rapid diagnostic technique was used to build computer-based learner-adapted learning environments (elementary algebra).


\[-3x = 7\]

Type in the first line of your solution and then press Return:

\[\text{-7/3}\]

[Don't know]
Solve for x:

\[
\frac{3x + 1}{3} = 1
\]

\[
\frac{3x + 1}{3} \times 3 = 1 \times 3
\]

\[
3x + 1 = 3
\]

\[
3x + 1 - 1 = 3 - 1
\]

\[
3x = 2
\]

\[
\frac{3x}{3} = \frac{2}{3}
\]

\[
x = \frac{2}{3}
\]

Continue
Solve for $x$: \[
\frac{10x + 15}{4} = 2
\]

\[
\frac{10x + 15}{4} \times 4 = 2 \times 4
\]

\[
10x + 15 = 8
\]

Complete the solution, type in your final answer, and press Return:
Further research

Limits of applicability:
science? technical knowledge? second language learning? ill-defined domains?

Psychometric characteristics:
sample size, multidimensional models?

Building learner-adapted environments
s.kalyuga@eaa.unsw.edu.au