

Title & Speaker

Perennial Issues and Contemporary Concerns about Assessing Mathematics Learning Outcomes at Secondary Level

WONG Khoon Yoong (Dr)

Maths & Maths Education

National Institute of Education, NTU

khoon.y.wong@gmail.com

<http://math.nie.edu.sg/kywong>

My Encounter with Assessment

- 1960s: as student, O-level, A-level, tertiary; paper-and-pencil questions; similar to current practices
- 1970s – 1980s: as teacher, set similar exam papers; team leader of Cambridge markers (Malaysia)
- 1980s – now: as maths educator, teach assessment courses for teachers, conduct research about assessment
- 1996: Talk and paper on traditional vs. reformed assessment
- 2002: Chaired an Assessment Symposium at EARCOME 2
- 2005: Maths Teacher Conference on Assessment
- 2008 – 2012: SMAPP (more later)
- Today: Look back on these changes, personal, not systematic review

Your Focus?

1. What can you learn from the Assessment “Reforms” in the past 50 years?
2. What changes do you plan to make about your assessment practice next term? next year?

Perennial & Contemporary

Many issues; only 4 aspects about cognitive domain:

Assessment Aspects	Perennial Issues	Contemporary Concerns
What to assess?	Routine skills	Authentic, real-life contexts, modelling
Delivery	Paper-and-pencil	IT-based
Purposes	Summative (AoL), diagnostic	Formative (AfL), AaL
Standard of performance	Norm-referenced (NRT)	Criterion-referenced (CRT); Ipsative

In-situ Reflection

What? Routine Skills

- Routine skills, modelled after public exams
- Hardly changed over the decades; main changes about units (imperial to SI), colonial contexts
- Timed tests; assess speed and accuracy
- Routine questions may be solved in non-routine ways; grading will allow for different methods

Variety of Questions

- Perennial issue: Vary types of questions to reduce regurgitation and test higher order thinking
- 1988: MP Dr Tan Cheng Bock

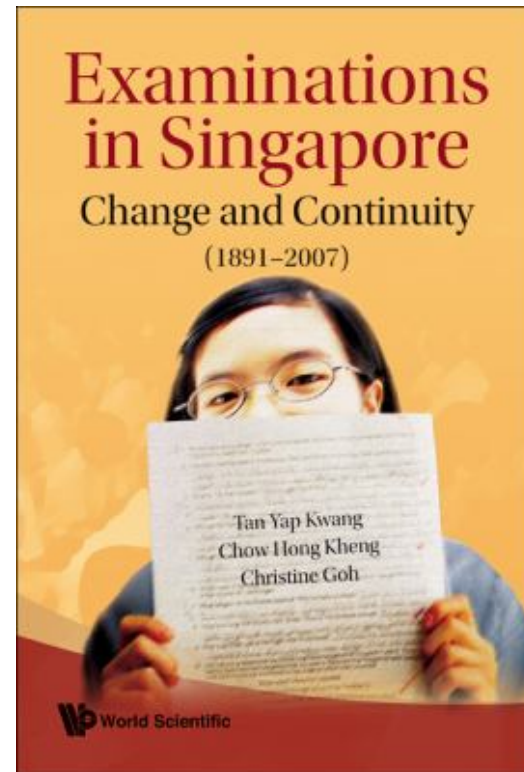


25/2/88
**Don't expect same
exams always: MP**

- Tay Eng Soon: There must be variations [in exams]. That element of **surprise** must be there.
- Teachers are still advised not to be stuck with past exam questions

Real-life Applications

- Recent focus on “real-life applications”
- Actually a perennial issue
- A question from (p. 17)



Junior Cambridge, 1892

- The French Railway Companies have recently reduced their rates for passengers as follows: 10 per cent reduction on the old first class fares, 20 per cent on the second class, and 30 per cent on the third class fares.
- In consequence of this reduction they anticipate an increase in the number of people travelling per mile as follows: an increase of 15 per cent in the number of first class passengers, 25 per cent in the second class, and 40 per cent in the third class.
- How much will the Railway Companies gain or lose per cent on each class?

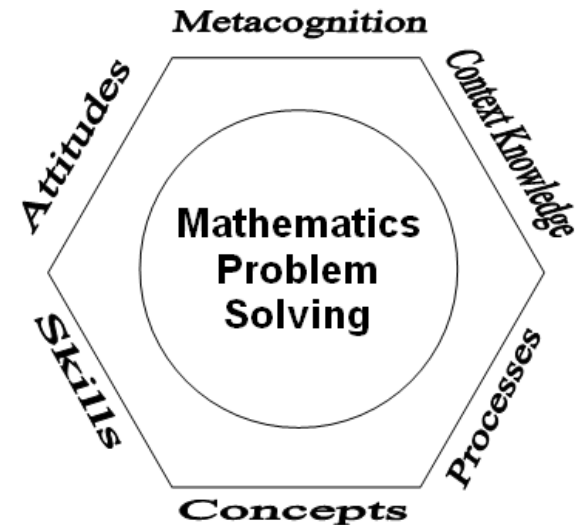
Junior Cambridge, Answer

Authentic & Alternative

- Mathematics Assessment Project, 2004 – 2007;
http://www.nie.edu.sg/files/oer/NIE_research_brief_10_003.pdf
 - Project assessment
 - Performance assessment
 - Student self-assessment
 - Communication assessment: journal writing, oral presentations
- Used in continual assessment

Knowledge about Contexts

- To solve *real* real-life problems, need knowledge about contexts, e.g., discount (pay less); fair dice
- Help students to acquire knowledge about contexts, e.g., loudness measured in decibel (dB); For every 3 dB over 85 dB, the exposure time before damage occurs is decreased by half.
- An enhanced curriculum framework



Delivery

- Traditional: Paper-and-pencil, timed
- Contemporary: IT-based, computer-adaptive testing (CAT), capture MCQ, fill-in-the-blank, graphing, constructed solutions
 - International trend; PISA IT-based problem solving
 - US: assessment consortia
 - SMAPP project

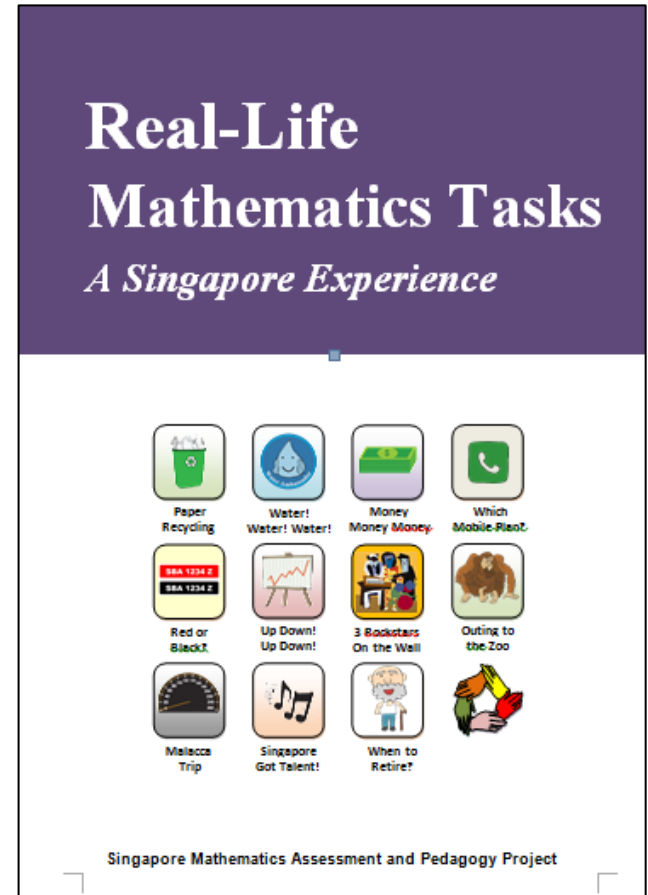
US Assessment Consortia

- To assess new curriculum, Common Core State Standards (CCSS)
- Partnership for Assessment of Readiness for Colleges and Careers (PARCC); try sample items for Grade 6-8 Maths (no answers) at: <http://practice.parcctestnav.com/#>
- Smarter Balanced Assessment Consortium (SBAC); need to log in

SMAPP

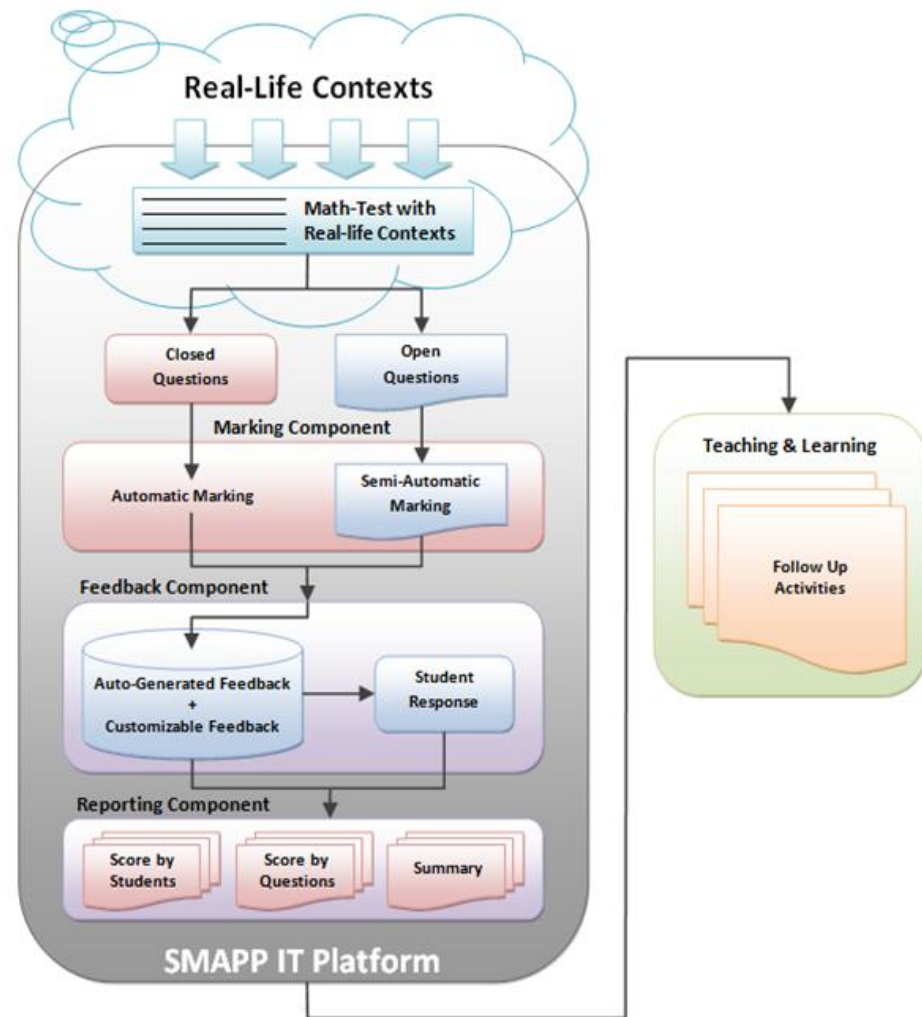
- Singapore Mathematics Assessment and Pedagogy Project; 2008 – 2012
- Real-life contexts, multiple maths competencies, AfL
- Print and electronic materials given to all secondary schools
- E-book:

<http://hdl.handle.net/10497/11492>



SMAPP IT Platform

1. **Deliver** extended tasks
2. **Capture** answers; closed and open-ended questions
3. **Automatic** marking of closed items; save time, consistent
4. **Semi-automatic marking** of open-ended items; consistency, aware of student thinking
5. **Customisable feedback**
6. Student responses to online feedback; rarely studied
7. Reports by questions, students, class; **feedback to teachers**
8. Follow up activities (non-IT)



SMAPP: Paper Recycling

- HTML version: Delivery and Capture

In-situ Reflection

Purposes

Types	Issues
Summative, Assessment <i>of</i> Learning (AoL)	<ul style="list-style-type: none">• Long standing purpose• A few times per year, end of term, year• Teaching to the test• Stressful?
Formative, Assessment <i>for</i> Learning (AfL)	<ul style="list-style-type: none">• Relatively recent• Frequent, mini-SA?• Feedback to students• Feedback to teachers• Diagnostic assessment
Assessment <i>as</i> Learning (AaL)	<ul style="list-style-type: none">• Student self-assessment

Summative: Stressful

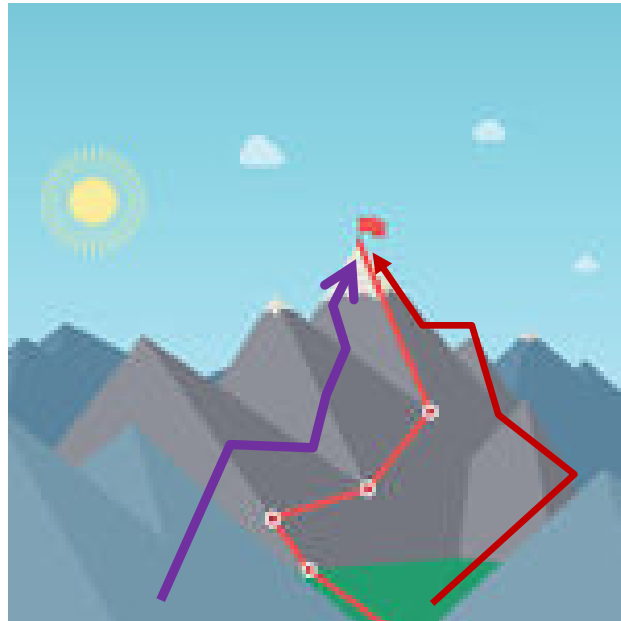
- School SA
- Public exams; stressful because results determine future education journey
- Pros: stress → motivation, prepare for life
- Cons: lack of sleep, poor health, tuition, lack of opportunities to explore “passions”
- NUS High: Results of Year 1 & 2 not “counted”
- Extend to all Secondary 1 classes in all schools?
Smoothen transition from Primary to Secondary level

Teaching to Test: Pros & Cons

- Common issue in many countries
- WYAIWYT (What You Assess Is What You Teach)
- Pro: If assessment covers high level contents, teaching will be enhanced
- Con: Assessment tends to cover low cognitive demand, teaching is restricted

Teaching to Test: Teaching

- Teach to-be-assessed contents directly is **not** necessarily the “best” route to success
- Different learning paths to achieve same goals; depend on student characteristics and learning experiences



Good (Smart) Students

- Achieving approach (Biggs)
- Their focus: tell me what I have to do to “score”
- May not care about whether their answers to problems make sense or not

Teaching to Test: Converse

- What You Teach Is What You Assess
- Agree or Disagree?

Teaching to Test: Converse

- What You Teach Is What You Assess
- Agree or Disagree?

Reporting: Changes

Issues	Examples
Perennial	<ul style="list-style-type: none">• School: Raw scores, percents of maximum scores, ranks• Public exams: Standardised scores; T-scores, letter grades
Contemporary	<ul style="list-style-type: none">• Performance categories or levels• Academic bands• PSLE; from aggregate T-scores to letter grades; potential impacts?• Profiles: strands by levels (not yet)

Failure Not Allowed?

- Some countries do not allow students to be graded as “fail”
- Western Australia (2006)

Old

- Exceptional
- Proficient
- Adequate
- Not adequate

“New”

- A: Excellent
 - B: Good
 - C: Satisfactory
 - D: Limited
 - E: Below minimal acceptable level
-

In-situ Reflection

Formative: History

- Modified from Cizek (2010)
- Scriven (1967); *formative evaluation* of curriculum
- Bloom et al. (1971); *summative vs. formative* evaluation of student learning
- Sadler (1989); gaps and feedback loop
- Black & Wiliam (1998); moderate effect sizes
- Wiliam (2007); 5 key strategies: criteria, evidence, feedback, resources, ownership
- Hattie (2009, 2012); teacher feedback, large effect size, 0.73, about 1.5 grade difference (e.g., C to B+)

Feedback: Types

- Feedback should be immediate and specific

Ways

- Moment-by-moment feedback during lessons; oral Q&A
- Observations of individual and group learning

Nature

- Homework, written comments, go over in class
- Extended projects, performance tasks, written comments
- Tests; marking, go over in class

Butler: Which Feedback?

- You return answer scripts to students.
- Which kind of feedback is helpful?
 - a) Comments only, e.g., You thought of quite a few ideas; maybe it is possible to think of more different ideas
 - b) Grades only
 - c) Comments and grades
- Your choice?
- Cited in Boaler (2009); Wiliam (2007)

Diagnostic Assessment

- What students are weak in by specific topics or skills
- Pre-conceptions: what they students bring into lessons from their life experiences
- Misconceptions: mistakes they have “learned”
- 1970s, 1980s; developed diagnostic in primary maths
- Contemporary: popular among Masters students, misconceptions in various topics, error analysis
- Less well-developed: diagnostic teaching, cf. Alan Bell (1992)

Same Item Different Purposes

- Solve $2x^2 - 3x - 2 = 0$ by factorisation.

Purposes	Nature
	Identify which aspect of the item (e.g., coefficient of x^2 term) is causing difficulty
	Use result to prepare follow-up lesson, feedback to student
	As part of student's attainment in the topic
	To determine whether the curriculum objective has been achieved, e.g., if students solve it by formula, then curriculum objective of factorisation has not been achieved

SMAPP: Standard Feedback

- Standard feedback to students:
 - Correct; general praise; highlight key steps; extend
 - Partially; correct steps, check key things
 - Wrong; give hints
 - Answer only; remind students to show working

F1 : Great work! It is

Select feedback

F0 : You seem to have skipped this question. Next time, try to answer all the questions

F1 : Great work! It is surprising to know that we can save so many trees just by using recycled paper!

F2 : You have 2 correct steps. Check the units.

F3 : You have 1 correct step. Check the other step.

F4 : First, find the total mass of paper that is used by the students. Second, convert the mass to tonnes. Third, use ratio or proportion to find the answer.

F99 : Next time, show your working.

SMAPP: Customisable Feedback

- Entered by teacher

F4 : 80g/meter square ▼

- ✖ 🔍

📊 | x_2 x^2 | + - \times \div / | = \neq \approx | $\sqrt{\quad}$ \angle \circ | \leq \geq | \pm

80 g/m² means 1 m² of the paper weighs 80 grams. The area of the A4 paper is Therefore it weighs.....|

Feedback to Teacher

- Hattie; effect size = 0.90
- How do teachers use data about student performance?
 - Those who do well
 - Misconceptions
 - Compare this performance with previous results
 - Class level; are curriculum goals achieved?

SMAPP: Volunteers

Volunteers needed for Trials
Interested? Contact Me

In-situ Reflection

Standards of Performance

- Different ways to make judgment about standards of student performance

Ways

Meaning

- Compare student performance with peers: current and/or past cohorts
- Public exams; reliability, psychometric
- Compare student performance to well-defined criteria (curriculum goals)
- Validity
- Compare student performance to previous performance; progress
- Used in physical performance; rare in academic work (re-do test?)

Ipsative?

- Solve $2x^2 - 3x - 2 = 0$ by factorisation.
- *Ipsative*: Re-do after a short lapse, average of first and second attempt, to determine whether student has improved

School vs. Life

Your Focus?

1. What have you learned from the Assessment “Reforms” in the past 50 years?
2. 1 = Nothing or a little; 5 = Some; 9 = A lot
3. What changes do you plan to make about your assessment practice next term? next year?
4. 0 = No change; 5 = Small change; 9 = Big change

Concluding Remarks
