Investigating the Intellectual Quality of

Alternative Assessments at Nanyang Girls' High School

Mrs Seah-Tay Hui Yong

Nanyang Girls' High School

Abstract

A unique feature of the Integrated Programme at NYGH is Student-Initiated Assessments (SIA). These are alternative assessments which generally involve authentic tasks that provide students with opportunities to integrate and apply classroom learning to the real world.

This paper will present the findings of an investigation into the intellectual quality of the SIA tasks set for the upper secondary Maths, Science and Language Arts classes using the scoring rubrics developed by Centre for Research in Pedagogy and Practice (CRPP). It will also discuss the implications of these findings within NYGH and beyond.

Background

In 2004, Nanyang Girls' High School embarked on the Integrated Programme which offers young, bright, and university-bound students the opportunity to skip the national examinations and proceed to Hwa Chong Institution for their "A" levels.

No longer constrained by the GCE "O" examination, the teachers designed an enriched and more challenging curriculum for the Integrated Programme. The format of assessment also moved away from the traditional paper-and-pen tests. Instead, students are encouraged to take responsibility for their own learning with the introduction of the Student-Initiated Assessment (SIA), which forms a considerable 30% of their overall grade. The SIA is a form of alternative assessment that generally

involves the use of authentic tasks that provide students with opportunities to practise higher order thinking skills, make connections between different domains of knowledge, and apply classroom learning to the real world situations. The rationale for implementing SIA is that a combination of traditional and alternative assessments will help obtain a richer and more complete picture of what students know and are able to do (Elliot, 1995).

Though relatively new to Singapore, such alternative assessments have been advocated in the United States and Australia as early as the 1950s. Bloom et al. had advocated that tasks such as "planning simple experiments" in elementary school science and "actually carrying out the experiments" can be "highly absorbing, more so than the usual run of school assignments" and "can challenge the student to do further work of a similar sort." (Bloom, et al., 1965, p. 167-168). More recently, alternative assessments have been advocated by educationists such as Wiggins (1993) and Newmann (2000) who variously refer to them as performance assessment (the accent being the ability to carry out a task) and authentic assessment (the accent being on "real world" tasks).

The premise underlying such alternative assessment is that while traditional tests do little more than require students to memorize information and give a fixed response, the former requires the student to apply that information to a particular situation, create a response and then explain or defend it. Such a process involves the use of higher-order thinking skills (e.g., cause and effect analysis, deductive or inductive reasoning, experimentation, and problem solving) (McBrien & Brandt, 1997).

The movement away from traditional tests was also the result of findings from contemporary cognitive psychology which indicates that learning is not just an accumulation of information but it is "an active process of mental construction and sense making." (Shepard, 2000, p. 6). Hence, "school learning should be authentic and connected to the world outside of school not only to make learning more interesting and motivating to students but also to develop the ability to use knowledge in real-world settings." (p. 7).

The promise of authentic assessment lies in its greater emphasis on metacognitive processes involved in problem-solving and higher-order thinking. However, unless such assessments are carefully designed, they may not end up even requiring the use of such of such complex cognitive processes. It cannot be assumed that they will do so by virtue of its real-life context.

The purpose of this study is to investigate whether the Student-Initiated Assessments at Nanyang Girls' High School yield valid information about desired authentic learning outcomes such as higher order thinking and deeper mastery of content knowledge.

Method

The written instructions and descriptions for upper secondary Language Arts, Maths and Science SIA tasks were collated and scored against rubric drawn up from the authentic intellectual quality standards designed by Panel 5 researchers of Centre for Research in Pedagogy and Practice (CRPP) Core Research Programme.

With their permission, I adapted their rubric to focus on the following 4 areas:

1. DEPTH OF KNOWLEDGE AND UNDERSTANDING

    1.1 Factual knowledge – grasp of basic knowledge

    1.2 Procedural Knowledge- grasp of basic procedures and processes

    1.3 Grasp of advanced concepts

2. KNOWLEDGE MANIPULATION

    2.1 Requires student to organize / interpret / analyze / synthesize / evaluate information/data

    2.2 Requires student to apply information gathered

    2.3 Requires student to generate or construct knowledge new to student; generate alternative perspectives / solutions

3. KNOWLEDGE CRITICISM

    3.1 Requires student to apply critical thinking to information rather than accept as given or truth

    3.2 Requires student to compare and contrast different sources of information or ideas

    3.3 Requires student to critique information / knowledge

4. CONNECTIONS TO THE REAL WORLD BEYOND CLASSROOM

I also sought the help of three other colleagues who had participated in the CRPP teacher-moderated judgment of student work like me in Nov 2004. Their help was crucial as firstly, one was a Science teacher; another, Maths; and lastly, Language Arts teacher like me. We all had experience using the CRPP authentic intellectual standards and scoring rubrics in our respective subjects during our

professional development/work attachment. Secondly, by having two scorers agree on a consensus score after scoring the tasks individually increased the inter-rater reliability. So I worked with each of the 3 colleagues to score the tasks in their respective subject area to arrive at consensus scores which are reported in Tables 1-4.

<u>Results</u>

One of the difficulties in this study was to find a common yardstick to judge the SIA tasks across the different subject areas. The "Task Scoring Rubric" proved to be a useful instrument - generic enough to be used meaningfully in the 3 very different subjects yet detailed enough to help identify 10 specific items that formed the criteria for good authentic assessment.

Table 1   *Summary of Consensus score of Upper Sec SIA tasks*

	Science n = 12		Maths n = 4		Language Arts n = 7	
	M	SD	M	SD	M	SD
Depth of Knowledge						
Factual Knowledge	3.4	0.9	2.5	1.0	2.9	0.5
Procedural Knowledge	2.7	0.9	2.3	1.3	3.7	0.5
Advanced Concepts	2.7	0.5	3.3	0.5	2.7	1.5
Knowledge Manipulation						
Organize / interpret / analyze / synthesize / evaluate information/data	3.8	0.5	3.5	0.6	3.7	0.6
Apply Info Gathered	3.3	1.0	2.8	1.5	3.9	0.0
Construct New Knowledge	2.8	1.1	3.0	0.8	2.6	0.5
Knowledge Criticism						
Critique Info	2.3	1.0	2	1.2	2.4	1.0
Compare / Contrast Info	2.3	1.0	2.8	0.5	2.3	0.5
Real World Connections	3.7	0.8	2.0	1.2	3.0	1.5

On a scale of 1-4[1], the tasks for all 3 subject areas scored moderately high on Knowledge Manipulation, particularly item 2.1 (Requires student to organize / interpret / analyze / synthesize data) but low on Knowledge Criticism, particularly item 3.1 (Requires student to apply critical thinking to information rather than accept as given truth).

All the SIA tasks had an average score of 2 and above. In fact, except for 2 Science projects (out of 12) and the Maths portfolio, all the other 20 SIA tasks scored closer to 3 and above; in general terms, there was some evidence that each of the tasks satisfied the criteria of good alternative assessments.

Details of each subject area are as follows.

(The conclusions drawn here have been refined after they were put to the HoDs of the respective subjects for critique and verification.)

In the Science SIAs, Real World Connection score was highest of the four components. This is as intended by the department as SIAs were meant to complement the traditional paper-pen tests (which rigorously test concepts) by providing students with opportunities to see how Science concepts play out in the real world.

This approach also explains why the Science SIA tasks scored relatively low on item 2.3 (Requires student to generate or construct knowledge new to student; generate alternative perspective / solution) and Knowledge Criticism (as mentioned earlier). For the same reason, the Science SIA tasks scored relatively low on item 1.2 (Grasp of basic processes) because procedural knowledge is tested in SPA (School-based Science Practical Assessment). In fact, SPA marks form part of a percentage of SIA marks.

---

[1] 1 – no evidence   2 – little evidence   3 – some evidence   4 – much evidence

Table 2. *Summary of Consensus score of Upper Sec Science SIA tasks*

			Item	Factual Knowledge 1.1	Procedural Knowledge 1.2	Advanced Concepts 1.3	Manipulate Data 2.1	Apply Info Gathered 2.2	Construct New Knowledge 2.3	Critique Info 3.1	Compare / Contrast Info 3.2	Real World Connections 4	Average score	Std Dev
Science	Biology	Sec 3	Designing a Multimedia Educational Package	4	2	2	4	3	2	3	2	4	2.9	0.9
		Sec 4	Analysing Journal Article	2	1	3	4	1	1	3	1	2	2.0	1.1
			Attending a Talk: pre/post activities	2	1	3	4	3	3	2	2	4	2.7	1.0
	Chemistry	Sec 3	Task 1: Separation Techniques to recover one natural product	4	3	2	4	4	1	2	3	4	3.0	1.1
			Task 2: Redox Reactions in everyday life	4	3	3	4	4	3	2	3	4	3.3	0.7
			Task 3: Investigating Food Chemistry	4	3	2	4	4	2	2	2	4	3.0	1.0
			Task 4: Designing Sabbatical on Chemistry Magic	4	3	3	4	4	4	2	1	4	3.2	1.1
		Sec 4	Jurong Island: pre/post visit activities	4	2	3	3	4	4	2	3	4	3.2	0.8
			Movie: Critical comment on accuracy on facts of science presented in movie	2	2	3	4	4	4	4	4	4	3.4	0.9
	Physics	Sec 3	Designing Optical toy / instrument	4	3	2	3	3	4	1	2	4	2.9	1.1
			Designing a Poster / Mindmap	3	1	3	4	2	2	1	3	2	2.3	1.0
		Sec 4	Review of a Physics-related Article	4	1	3	3	3	3	4	1	4	2.9	1.2
			Component Score		2.7			3.3			2.3	3.7	2.9	0.6

Unlike the other two subjects, Maths SIA tasks scored low on Real World Connection.

However, the reason lies in the intention of the Maths department which was to offer to students a variety of tasks, hence differentiating to cater to different students' interest to complement traditional paper-pen tests (which are designed to assess students' grasp of concepts). It is worth noting that depth is offered through the Australian Maths Enrichment Series which scored high on Depth of Knowledge and Knowledge Manipulation.

There is also a case for Maths at this level to be more esoteric, dealing with theorems and proofs; hence the low score on Real World Connection.

Table 3. *Summary of Consensus score of Upper Sec Maths SIA tasks*

		Item	Factual Knowledge 1.1	Procedural Knowledge 1.2	Advanced Concepts 1.3	Manipulate Data 2.1	Apply Info Gathered 2.2	Construct New Knowledge 2.3	Critique Info 3.1	Compare / Contrast Info 3.2	Real World Connections 4	Average score	Std Dev
Maths	Sec 3 & 4	Portfolio	2	1	3	3	1	3	1	3	1	2.0	1.0
		Australian Maths Enrichment Series	4	4	4	4	4	4	1	2	1	3.1	1.4
		Project Work	2	2	3	4	4	3	3	3	3	3.0	0.7
		Independent Learning: research into topic of student's choice	2	2	3	3	2	2	3	3	3	2.6	0.5
		Component Score	2.7			3.1			2.4			2.3	0.5

The Language Arts Programme in our school focuses on equipping students with skills so that they are able to communicate effectively for academic, literary and functional purposes. The SIAs were conceived as opportunities to serve mainly the functional strand, especially oral competency which accounts for one third of SIA mark. It is thus no surprise when the tasks generally scored high on item 4 (Real world application) as mentioned before, and items 1.2 ("Procedural knowledge") and 2.2 ("Apply information gathered").

Of the four components, Knowledge Criticism was the only that did not score above 3.

Table 4. *Summary of Consensus score of Upper Sec language Arts SIA tasks*

			Item	Factual Knowledge 1.1	Procedural Knowledge 1.2	Advanced Concepts 1.3	Manipulate Data 2.1	Apply Info Gathered 2.2	Construct New Knowledge 2.3	Critique Info 3.1	Compare / Contrast Info 3.2	Real World Connections 4	Average score	Std Dev
Language Arts	Sec 3	Oral Tasks	Passionata Speech	3	3	3	4	4	2	1	1	4	2.8	1.2
			Socratic Seminar	4	4	4	4	4	4	4	4	4	4.0	0.0
	Sec 4		Presentation of topic (student's choice)	2	4	3	4	3	1	2	2	2	2.6	1.0
			Speech: Excerpt from Shakespeare	3	4	4	4	4	3	1	2	1	2.9	1.3

	Sec 3	Organising an Event		3	4	1	4	4	3	3	2	4	3.1	1.1
		Letter to the Editor		3	4	1	3	4	3	3	3	4	3.1	0.9
	Sec 4	Learning Portfolio of writing		2	3	3	3	4	2	3	2	2	2.7	0.7
			Component Score	3.1			3.4			2.4			3.0	0.6

Conclusion

Many have confirmed the wisdom in the old "WYTIWYG" adage – "What you test is what you get". When teachers set assessments that demand in-depth knowledge and higher-order skills, students have risen to the occasion to produce higher quality intellectual work. A major study of student work by Newmann and his colleagues found that when teachers set "authentic" intellectual work - work that required high-level cognitive performance as well value beyond school- students tended to perform at higher levels (Newmann, Lopez and Bryk, 1998). In a case study of 5 schools which developed "authentic" performance-based assessment, it was reported that there was widespread evidence of in-depth learning, intellectual habits of mind, quality products (Darling-Hammond, Ancess, & Falk, 1993). The students involved in their research were typically those who scored low on traditional, standardized assessment, from low-income and working class families. Yet they were able to master complex content and higher order skills, something they were not able to do so previously because presumably they were denied the opportunity.

If this is the case, SIAs are promising instructional and assessment tools to develop in these already bright students the critical thinking and real-world problem solving skills needed to succeed in the 21st century.

However, the results underscore the need for all 3 subject areas to consider raising demands of Knowledge Criticism in the SIA tasks. This is probably unfamiliar

territory for teachers as GCE "O" level assessments have not required this of students. However, this should change in view of designing a more challenging curriculum for these more able students. Besides, in an era when change is the order of the day, we need to prepare students to exercise judgement and innovation to solve unstructured problems. One way is to require students to apply critical thinking to information rather than accept it as given truth.

References

Bloom, B.S., Engelhart, M.D., Furst, E.J., Hill, W.H., & Krathwohl, D.R. (Eds.)(1956). *Taxonomy of Educational Objectives. The Classification of Educational Goals, Handbook I: Cognitive Domain*. New York: David McKay Company, Inc.

Darling-Hammond, L., Ancess, J., & Falk, B. (1993). *Authentic assessment in action: studies of schools and students at work.* New York: Teachers College Press.

Elliot, S. N. (1995). Creating Meaningful Performance Assessments: ERIC Digest E531. *ERIC Clearinghouse on Disabilities and Gifted Education Reston VA*. Retrieved July 4, 2005, from http://www.ericdigests.org/1996-1/creating.htm

McBrien, J. L. & Brandt, R. S., (1997). *The Language of Learning: A Guide to Education Terms.* Retrieved May 2, 2005, from http://www.ascd.org/portal/site/ascd/menuitem.4427471c9d076 deddeb3ffdb62108a0c/

Newmann, F. M. (2000). Authentic intellectual work: What and why? *Research / Practice, 8(1).* Retrieved May, 2, 2005, from http://education.umn.edu/CAREI/ Reports/Rpractice/Fall2000/newmann.html

Newmann, F. M., Lopez, G. & Bryk, A.S. (1998). *The quality of intellectual work in Chicago schools: A baseline report.* Chicago: Consortium on Chicago School Research. Retrieved February 25, 2006, from http://www.consortium-chicago.org/publications/p0f04.html

Shepard, L. (2000). The role of assessment in a learning culture. *Educational Researcher, 29* (7), 4-14. Retrieved March, 6, 2006 from JSTOR database.

Wiggins, G. P. (1993). *Assessing student performance: Exploring the purpose and limits of testing.* San Francisco: Jossey-Bass Publishers.