

**Development of the Singapore Prototype Classroom Assessment Tasks:
Innovative Tools for Improving Student Learning and Performance**

Kim Koh, Ai Noi Lee, Wengao Gong, Hwei Ming Wong

**Centre for Research in Pedagogy and Practice, National Institute of Education
Nanyang Technological University**

Abstract

This paper describes the development of the Singapore prototype classroom assessment tasks in English, Mathematics, Science, Chinese Language, Malay Language, and Tamil Language as well as the use of authentic intellectual standards and scoring rubrics for evaluating the quality of assessment tasks and related students' work. The authors will provide a brief background to and the rationale for the development of standards for assessing the quality of classroom assessment tasks and related students' work across subject areas at Primary grade levels in Singapore schools. This will be followed by in-depth discussions of the alignment of the prototype assessment tasks with the authentic intellectual standards; the training of teachers in crafting high authentic intellectual assessment tasks; the implementation of the assessment tasks in the day-to-day classroom teaching, learning, and assessment; the construction of the scoring rubrics and exemplars; the moderation procedures for teachers to evaluate the quality of students' work in response to the authentic intellectual assessment tasks; and the strategies in empowering teachers to build professional learning communities within and between schools. Findings and implications of the prototype classroom assessment tasks on teachers' pedagogical and assessment practices as well as on student learning and performance will also be discussed.

Introduction

In Singapore, there are many important initiatives launched by the Ministry of Education to reform the nation's educational system. For example, 'Thinking Schools, Learning Nation'; 'Innovation and Enterprise'; 'Teach Less Learn More'; and 'Engage Learners' are the government's visions for developing a productive and resilient nation to face the challenges in the knowledge-based economy. All these visions have led to the implementation of teaching for higher-order thinking skills rather than for rote knowledge. Singaporean teachers are encouraged to move away from the conventional didactic teaching methods to constructivist learning approaches in order to promote students' higher-order thinking skills, problem-solving skills, and communication skills. Likewise, assessment approaches have gone beyond assessing students' knowledge of facts and mastery of basic skills. Many educators and educational researchers believe that traditional paper-and-pencil testing (e.g., multiple-choice items, fill-in-the-blank questions) is limited to assessing students' lower-order thinking skills. Hence, it should be gradually replaced or be at least complemented by alternative or authentic assessments such as project work, portfolios, major research paper, performance-based assessment (e.g., SPA¹), and student self-assessment (e.g., SAIL²). The aforementioned types of assessment allow for assessing students' higher-order thinking skills and real-world problem solving. When well designed and properly implemented, alternative assessments are closely aligned with curriculum and instruction that emphasize knowledge construction and problem solving in real-world contexts (Herman, Aschbacher, & Winters, 1992). This is in line with Shepard's (2000) argument for a new "social constructivist" paradigm for classroom learning and assessment. She contended that assessment should be an integral part of ongoing instruction aimed at the development of critical thinking, problem solving, application, and metacognition.

Wiggins (1989) suggests that a truly authentic/innovative assessment system should meet the following five criteria: (1) criterion-/standards-referenced, (2) formative, (3) moderated, (4) clear in the progression of educational development (e.g., score levels on a rubric would reflect the criteria or standards), and (5) a substantive assessment framework that describes the achievement variables that are valued and thus worth assessing. According to Wilson (1994), local teacher moderation is an example of a community of judgment that functions as a motivation for teacher change as well as a catalyst for changing the assessment culture of the school or district. Teacher moderation gives teachers ample opportunities to reflect on their own assessment practices, to interact with other teachers, to discuss instructional implications as well as to agree upon standards for authentic intellectual assessment. Furthermore, moderation affects the norms of the professional culture within a school or district, which is essential for "changing norms of practice and pedagogy" (Loucks-Horsley, Hewson, Love, & Stiles, 1998, p.105). The involvement of teachers in local assessment moderation serves as a medium for high-quality professional development. It adds significantly to teachers' skills for assessing student learning and improving their own teaching. Hargreaves has pointed out, "high-quality professional development for teachers is indispensable to

¹ SPA = School-based Science Practical Assessment.

² SAIL = Strategies for Active and Independent Learning.

bringing about deep and lasting changes in pupil achievement” (Hargreaves, 2003, p. 126).

Previous research on the quality of classroom assessment and student work includes the work of Bryk, Nagaoka, and Newmann (2000); Lingard and Ladwig (2001); Luke, Matters, Herschell, Grace, Barrett, and Land (2000); Newmann and associates (1996a), Newmann, Marks, & Gamoran, (1996b), to name just a few. In the Newmann et al. (1996a & 1996b) and Bryk et al. (2000) studies of Chicago teachers’ assignments in mathematics and writing in grades 3, 6, and 8, trend analyses of student learning gains on basic skills, as measured by standardized tests, were conducted to examine academic progress in the selected schools. An in-depth longitudinal study (1997-1999) of the quality of intellectual work was carried out by the researchers in order to complement the test score analyses. They found that students who received assignments requiring more challenging intellectual work achieved greater than average gains on the Iowa Tests of Basic Skills in reading and mathematics, and demonstrated higher performance in reading, mathematics, and writing on Illinois Goals Assessment Program. Newmann et al. have also found high quality assignments in some very disadvantaged Chicago classrooms and that all students in these classes benefited from exposure to such instruction. There was a strong relationship between the quality of teacher assignments and student work. When teachers assigned more intellectually demanding classroom tasks, most students were able to demonstrate more complex intellectual performance. Similar to Newmann et al., the Lingard and Ladwig (2001) and Luke et al. (2000) study found that Australian students’ performance were strongly dependent upon what was asked of them in the assessment tasks. When teachers focused on high-level performance and had high expectations of their students, students regardless of their backgrounds were more likely to achieve greater intellectual quality in their classroom work and basic skills.

Based on the CRPP initial findings of Panel 5 research, we found that teachers in all subjects except Primary 5 Social Studies tended to place more emphasis on drill and practice of basic knowledge and skills. The results of the quality of teacher assessment tasks suggest that the majority of the assignments required low levels of understanding of subject matter knowledge and higher-order thinking. In general, students were asked to regurgitate information given by the teachers or textbooks, to reproduce memorized basic concepts or algorithms, or to repeat use of previously learned skills or procedures. They were not given enough opportunities to involve in the learning and mastery of advanced concepts and to engage in higher-order thinking nor real-world problem solving. Although teachers had required sustained writing in the assignments, especially in English and the quality of student writing was scored as moderately high, it does not suggest that students in this study were able to communicate effectively in elaborative forms. This is because most of the essays were in the forms of simple narrative and description. Furthermore, the contents, sentence structures, and organization were almost similar.

The results of the quality of student work echo the message “what you test is what you get”. The majority of the student work in all subjects except Primary 5 Social

Studies demonstrated high-levels of regurgitation and reproduction of factual knowledge and procedural knowledge. When teachers' assessment tasks did not require high authentic intellectual work, most students were not able to demonstrate complex intellectual performance.

In the Panel 5's 2004/05 teacher moderation sessions, many participating teachers have expressed the needs for an in-service professional development program conducted by NIE/CRPP in improving Singaporean teachers' assessment literacy, especially their capacities in designing high authentic intellectual assessment tasks that elicit students' higher-order thinking and real-world problem solving. Teacher assessment literacy includes knowledge and skills in learning and assessment, classroom assessment task design, performance standards and scoring rubrics.

Research on teachers' assessment practices has shown that many teachers are inadequately trained and ill-prepared to develop, administer, and interpret the results of various types of assessments, including alternative assessments (Bol et al., 1998; Daniel & King, 1998; Plake, Impara, & Fager, 1993; Stiggins, 1992). Bol et al. found that the frequency with which teachers used different methods of assessment was related to their perceptions of how well prepared they were to develop and administer the assessments. Teachers who were less prepared and skilled in developing open-ended assessment tasks as well as in developing alternative assessments in general, perceived them to be more difficult to develop than traditional paper-and-pencil measures. Moreover, teachers' assessment practices were not aligned with their instructional goals and tended to demand low level of cognitive processing. Many teachers were found to be not good judges of the quality of their assessment tasks (Bol and Strage, 1996). For example, they thought the assessment tasks were assessing higher-order cognitive skills but this is not the case – the assessment tasks were actually measuring lower-order cognitive processes. This indicates that low assessment literacy might result in negative perception toward the use of alternative assessment. This could, in turn, adversely affect the quality of their assessment tasks. Two other variables that are worth examining: teaching experience and attitude toward scoring rubric. Researchers have found that teachers with the most teaching experience felt more prepared than teachers with the least experience to develop and administer alternative assessment (Bol et al., 1998). Apart from teaching experience, teachers' attitude toward the rubric used to score the open-ended tasks could play an important role in changing their assessment practices. Some teachers perceived that the use of scoring rubric required extra work. This hindered them from using alternative assessment in their day-to-day classroom instruction.

Objectives of Inquiry

In many places of the world, teachers are often left to fend for themselves in developing assessments for classroom use. Prior to the implementation of large-scale authentic assessment in Singapore classrooms, a set of well-designed and technically sound prototype assessment tasks must be made available to the teachers. This set of tasks will enable frequent and in-depth assessment of student understanding during the course of instruction. The prototype assessment tasks are deemed to provide Singaporean

teachers with a repertoire of assessment tools that can be used to diagnose and improve their students' learning.

The study aims to enhance teacher capacities in designing high-quality classroom assessment tasks and in using reliable and valid scoring rubrics to look at their students' work based on the CRPP authentic intellectual standards (Koh, Lee, Tan, & Wong, 2004). In essence, the scoring rubrics of assessment and student work as well as teacher moderation could serve as a heuristic for designing classroom assessment tasks aligned with the authentic intellectual standards, as a structure to help guide discussion of assessment tasks and student work, as a way for teachers to adopt the authentic intellectual standards in their teaching and testing for higher-order thinking as well as a tool for teachers' professional development. In addition, we aim to empower Singaporean teachers to build their professional learning communities within and between schools. This will enhance teachers' collective intelligence and social capital – including ways of sharing and developing knowledge and skills in alternative assessment – among fellow teachers (Hargreaves, 2003).

Methodology

Sample

This study consists of 16 Primary schools (8 experimental schools and 8 control schools) in Singapore. Our focus of assessment literacy intervention involves teachers who teach upper primary classes in the neighbourhood schools. The school samples are volunteer samples, depending on the interest of participation indicated by the schools' principals and teachers. This is because the willingness of principals and teachers to participate in the intervention will effect the teachers' commitment to the professional development sessions. The same cohorts of teachers and students will be followed from primary 4 (in year 2006) to primary 6 (in year 2008). The subject areas include English, Mathematics, Science, Chinese Language, Malay Language, and Tamil Language.

Authentic Intellectual Quality of Assessment Tasks and Student Work

Two sets of standards are used to evaluate the authentic intellectual quality of teacher assessment tasks and student work (see Table 1). These standards will also be used to train and guide teachers in designing high authentic intellectual assessment tasks.

Table 1. *Standards for Evaluating the Quality of Teacher Assessment Tasks and Student Work*

Standards for Teacher Assessment Tasks	Standards for Student Work
Depth of Knowledge	Depth of Knowledge
Knowledge Criticism	Knowledge Criticism
Knowledge Manipulation	Knowledge Manipulation
Sustained Writing	Sustained Writing
Text Types	Text Types
Clarity and Organization	Quality of Student Writing /Answers
Connections to the Real World beyond the	Connections to the Real World

Classroom
Supportive Task Framing
Student Control

beyond the Classroom

Standards for Teacher Assessment Tasks

Under *depth of knowledge*, we conceptualized three forms of knowledge: factual knowledge, procedural knowledge, and advanced concepts according to revised Bloom's taxonomy of knowledge (Anderson & Krathwohl, 2001). We have particular interest in the degree to which a teacher has demanded students to perform each form of the knowledge in the classroom assessment tasks. Higher-order thinking is represented by *knowledge criticism* and *knowledge manipulation*. *Knowledge criticism* is exemplified by tasks that ask students to compare and contrast knowledge and to critique knowledge whereas *knowledge manipulation* is exemplified by tasks that demand students to manipulate or construct information and ideas. *Knowledge manipulation* includes organizing, analyzing, interpreting, synthesizing, or evaluating information; applying information to solve new problems; or arriving at new conclusions that produce new meanings or understandings for them (Newmann, Marks, & Gamoran, 1996). Two negatively-worded standards, namely *presentation of knowledge as given* and *reproduction*, were purposely included to enhance the reliability of scoring. High scores on these two standards indicate low levels of task demands for higher-order thinking.

Sustained writing refers to tasks that demand students to demonstrate their ability to elaborate on their answers through extended narratives, explanations, and justifications. In Mathematics, the use of prose, tables, equations, or diagrams is subsumed under sustained writing. *Making connections to the real world beyond the classroom* is evidenced by tasks that required students to apply knowledge and skills to the real world problems or to make connections between knowledge and personal meanings. *Text types* were adopted from the PISA assessment (OECD, 2000). Knowledge of *text-types* is considered as one of the indicators of deep understanding of content knowledge (Aschbacher, 1999).

We contend that teacher's *supportive task framing* will result in higher authentic intellectual quality in student work. In alternative assessment, teacher's scaffolding of an assessment task {i.e., written or graphic guidance and structure that assist students to complete a task is a desirable element (Nitko, 2004)}. *Task clarity and organization* is another important standard that provides positive framing for students to understand and meet the task requirements (Hunter, 1982). *Student control* and *explicit performance standards/ marking criteria* are indicators of learner support (Marzano, 1992).

Standards for Student Work

Our findings have shown that 'what you test is what you get', i.e., teachers' task demands for high authentic intellectual work will lead to high authentic intellectual quality of students' produced work. Hence, similar standards were used to evaluate the quality of student work. Students' understanding of factual knowledge, procedural knowledge, and advanced concepts are subsumed under *depth of knowledge*. Higher-order thinking is reflected by the ability to critique and manipulate knowledge

(*knowledge criticism and knowledge manipulation*). The theoretical underpinnings for *making connections to the real world beyond the classroom, sustained writing, and text types* are similar to those of the assessment tasks. *Quality of student writing/ answers* is demonstrated by clear, accurate, and convincing answers.

Teacher Professional Development Sessions

In the teacher professional development sessions, assessment task design, the use of standards and scoring rubrics, and teacher moderation of actual assessment tasks and student work from their schools will be conducted. Only teachers coming from the experimental schools are involved in the professional development. ‘Looking Together at Student Work’ teacher professional development model developed by Blythe, Allen, and Powell (1999) from the Teachers College, Columbia University will be modified for use with the Singaporean teachers. Teacher professional development sessions will be conducted during the school holidays in each academic year. CRPP researchers will act as teachers’ collaborators or critical friends in their classroom instruction. This ensures that teachers in the experimental schools will implement the authentic intellectual standards in their classroom assessment.

Prototype Assessment Tasks

Example: Primary 4 Chinese Language

单元 2-2 来自美国的 电邮	写电邮：假设你是妹妹，请你回一封电邮给姐姐，告诉她父母亲如何为你庆祝生日。也把你在生日当天所拍的照片传送给她。	<p><i>Authentic Intellectual Quality:</i> Integrating (making meaningful connection between new information and prior knowledge), Organising (classifying, analyzing, communicating), peer interaction.</p> <p>Sustain writing, Text-type: Recount Core values.</p>
	做连环画：用连环画的形式（有文字，有图片），以《我的十岁生日》为题写一本小书，然后和你的老师和同学分享。作品的形式可以是手写的，也可以是电子的（电子书）。	<p><i>Authentic Intellectual Quality:</i> Integrating (making meaningful connection between new information and prior knowledge), Organizing (classifying, analyzing, communicating)</p>

		Sustain writing, Text-type: Recount Core values, IT skills
--	--	--

Due to space limitation, we only include an example of the prototype high authentic intellectual quality assessment tasks in primary 4 Chinese Language. The authentic intellectual standards are aligned to the primary 4 Chinese Language syllabus. The assessment tasks are shared with the teachers in the professional development monthly meetings. Teachers' feedback and needs are taken into account in developing the assessment tasks. After the sharing, the teachers are asked to incorporate the assessment tasks into their day-to-day curriculum for two terms. We hope that the assessment information will help teachers make appropriate decisions about their instructional methods that can assist learning. At the same time, the information will enable teachers to understand the nature of student learning and to provide better feedback to students, which in turn can enhance learning. Teacher assessment literacy is expected to improve after the end of the academic year. The prototype assessment tasks in other subjects will be presented at the conference.

Educational Implications

In Singapore, teachers have done excellently in delivering knowledge and skills that are essential for high-stakes norm-referenced standardized testing. This is evidenced by students' top performance on the international achievement assessments (e.g., TIMSS). However, it is timely for teachers to move toward more authentic type of assessment so that students can engage in higher-order thinking and real world problem solving. Conventional high-stakes tests can only assess students' discrete bits of factual and procedural knowledge by rote memorization. The learning process is detached from real life problems, which are often ill-structured. Due to the pressure of teaching to the tests, most teachers adhere to assessment formats and scoring practices found in high-stakes exams. Hence, it is important for teacher education programs and professional development workshops to equip teachers with contemporary knowledge about learning and assessment, especially the knowledge needed to develop high authentic intellectual tasks that would elicit students' complex thinking skills and make it possible to assess their growth and progress toward a set of high authentic intellectual standards (Cizek, 2000). The findings of this study will have an impact on the curriculum design of the teacher education programs in Singapore.

References

- Anderson, L.W., & Krathwohl, D.R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Longman.
- Aschbacher, P.R. (1999). *Developing indicators of classroom practice to monitor and support school reform, CSE Technical Report 513*. UCLA: National Center for Research on Evaluation, Standards, and Student Testing.
- Blythe, T., Allen, D., & Powell, B.S. (1999). *Looking together at student work: A companion guide to assessing student learning*. New York: Teachers College Press.
- Bol, L., Stephenson, P.L., O'Connell, A.A., & Nunnery, J.A. (1998). Influence of experience, grade level, and subject area on teachers' assessment practices. *The Journal of Educational Research*, 91(6), 323-330.
- Bol, L., & Strage, A. (1996). The contradiction between teachers' instructional goals and their assessment practices in high school biology courses. *Science Education*, 80(2), 145-163.
- Bryk, A.S., Nagaoka, J.K., & Newmann, F.M. (2000). *Chicago classroom demands for authentic intellectual work: Trends from 1997-1999*. Chicago: Consortium on Chicago School Research.
- Cizek, G.J. (2000). Pockets of resistance in the education revolution. *Educational Measurement: Issues and Practice*, 19(1), 16-23.
- Daniel, L.G., & King, D.A. (1998). Knowledge and use of testing and measurement literacy of elementary and secondary teachers. *Journal of Educational Research*, 91(6), 331-344.
- Hargreaves, A. (2003). *Teaching in the knowledge society*. Maidenhead: Open University Press.
- Herman, J.L., Aschbacher, P.R., & Winters, L. (1992). *A practical guide to alternative assessment*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Hunter, M. (1982). *Mastery teaching*. El Segundo, CA: TIP Publications.
- Koh, K., Lee, A.N., Tan, W., & Wong, H.M. (2004). *Manuals for scoring teacher assessment tasks and student work*. National Institute of Education: CRPP.
- Lingard, B., & Ladwig, J. (2001). *The Queensland School Reform Longitudinal Study: Teachers' summary*. Brisbane: Education Queensland.

- Loucks-Horsley, S., Hewson, P.W., Love, N., & Stiles, K.E. (1998). *Designing professional development for teachers of science and mathematics*. Thousand Oaks, CA: Corwin Press.
- Luke, A., Matters, G., Herschell, P., Grace, N., Barrett, R., & Land, R. (2000). *New Basics Project Technical Report*. Brisbane: Education Queensland.
- Marzano, R.J. (1992). *A different kind of classroom: Teaching with dimensions of learning*. Alexandria, VA: The Association for Supervision and Curriculum Development.
- Newmann, F.M., & Associates. (1996a). *Authentic achievement: Restructuring schools for intellectual quality*. San Francisco: Josey Bass.
- Newmann, FM, Marks, HM, & Gamoran, A. (1996b). Authentic pedagogy and student performance. *American Journal of Education*, 104 (8), 280-312.
- Nitko, A.J. (2004). *Educational assessment of students*. 4th Ed. Upper Saddle River, NJ: Pearson/Merrill Prentice Hall.
- Organization for Economic Co-operation and Development (2000). *Measuring student knowledge and skills: The PISA assessment framework of reading, mathematical and scientific literacy*. Paris: OECD.
- Plake, B.S., Impara, J.C., & Fager, J.J. (1993). Assessment competencies of teachers: A national survey. *Educational Measurement: Issues and Practice*, Winter, 10-12.
- Shepard, L. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4-14.
- Stiggins, R.J. (1992). *In teachers' hands: Investigating the practices of classroom assessment*. Albany: State University of New York Press.
- Wiggins, G. (1989). A true test: Toward more authentic and equitable assessment. *Phi Delta Kappan*, 703-713.
- Wilson, M. (1994). Community of judgment: A teacher-centered approach to educational accountability. In Office of Technology Assessment (Ed.), *Issues in educational accountability*. Washington, D.C.: Office of Technology Assessment, United States Congress.