Looking Collaboratively at the Quality of Teacher Assignments and Student Work in Singapore Schools

Kim Koh & Ai Noi Lee
National Institute of Education
Nanyang Technological University

Abstract

This study examined the quality of teacher assignments and associated student work in Singapore schools. Using the authentic intellectual quality framework, two sets of standards and scoring rubrics were developed for the training of teachers to judge the quality of assignments and student work. The samples of teacher assignments and student work were collected in English, Social Studies, Mathematics, and Science subject areas from 30 elementary schools and 29 high schools. There were significant differences for the authentic intellectual quality of teachers' assignments by subject area, stream, and grade level. Subject area effect was found to be larger than stream and grade level effects. Likewise, the differences of authentic intellectual quality for student work were significant and varied by subject area, stream, and grade level. Subject area effect was large. The correlations between the quality of teachers' assignment tasks and student work were strong and significant at both grade levels. The findings suggest that teacher professional development in high authentic intellectual quality task design is necessary for improving student learning and performance.

Please send all correspondence to the first author at khkoh@nie.edu.sg
Kim H. Koh, Ph.D.
Centre for Research in Pedagogy and Practice
National Institute of Education
Nanyang Technological University
1 Nanyang Walk, SINGAPORE 637616

Paper presented at the annual meeting of the American Educational Research Association, San Francisco, April 10, 2006.

Looking Collaboratively at the Quality of Teacher Assignments and Student Work in Singapore Schools

Objectives of Inquiry

Conventional assessment of student achievement has focused mainly on a high level of reproduction of factual and procedural knowledge from students. The items on such assessment typically measure recall of discrete facts, retrieval of given information, and application of routine computational formulas or procedures (Newmann, Lopez, Bryk, 1998). Often, the 'snapshot' assessment results give only a partial picture of students' intellectual performance at a given moment, but they are used as an indicator of teachers' job performance. To make sure that they are not penalized by low student performance, some teachers resolve to "teach to the test" by matching their instruction and assessment to the high-stakes exam formats rather than to the intended learning outcomes. Research on human learning and performance has documented that conventional assessments fail to measure students' higher order cognitive abilities or to support their capacities to perform real world tasks (Resnick, 1987).

In order to prepare students to become critical thinkers, productive workers, and lifelong learners in the knowledge-based economy, there is an urgent need for classroom assessment to move toward constructivist learning approaches to promote students' higher-order thinking skills, in-depth conceptual understanding, problem-solving skills, and communication skills (Darling-Hammond & Falk, 1997; Newmann & Associates, 1996; Shepard, 1989). It is important for teachers to design classroom assignments or assessment tasks that require students to demonstrate these authentic intellectual skills.

Previous research has shown that when teachers assigned more intellectually demanding assignments, students were able to demonstrate more complex intellectual performance in their work. Newmann and associates' (1996, 1998, & 2001) and Bryk, Nagaoka, and Newmann's (2000) studies examined the intellectual quality of teachers' assignments in mathematics and writing at grades 3, 6, and 8 in Chicago schools. 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. In addition, there was a strong relationship between the quality of teacher assignments and student work, i.e., teachers who assigned more intellectually demanding tasks were more likely to get authentic intellectual work from students. Similarly, Lingard and Ladwig (2001), and Luke, Matters, Herschell, Grace, Barrett, and Land (2000) have found that students' performance in Australia schools was dependent upon what was asked of them in the teachers' assessment tasks. In the Clare and Aschbacher (2001) and the Matsumura (2003) studies, the quality of the teacher assignments was found to be statistically significantly associated with the quality of classroom instruction and the quality of student work in language arts.

In Singapore, students' test scores from conventional assessments remain a key indicator of teachers' job performance and school effectiveness. However, the policymakers have realized that the teaching of thinking skills to Singaporean students is of paramount importance, especially in the knowledge-based economy. The importance of cultivating critical and creative thinking as well as innovative and lifelong learning has led to a systematic infusion of higher-order thinking skills into the nation' curriculum. Singaporean teachers are being encouraged to incorporate alternative assessments (e.g., project work, performance-based tasks, student self-assessment) into their classroom practices. A prevailing assumption underlying alternative assessments is that they encourage instructional strategies that foster reasoning, problem solving, and communication (Frederiksen & Colins, 1989; National Council on Education Standards and Testing, 1992).

To date, little information has been available to policymakers, school officials, and teachers regarding the classroom assessment practices and their effects on students' learning. This study aims to investigate the extent to which Singaporean teachers make authentic intellectual demands on students in their classroom assessment practices. The specific objectives of the inquiry are (a) to describe the patterns of the classroom assessment practices in both elementary and high schools, (b) to examine the quality of teacher assignments or assessment tasks, (c) to examine the quality of student work in response to the teacher assignments or assessment tasks, and (d) to determine the relationship between the quality of teacher assignments and the quality of student work.

Theoretical Framework

Newmann et al.'s (1996) "authentic intellectual work" consists of three standards: construction of knowledge, disciplined inquiry, and value beyond school. Authentic intellectual work enables students to engage in higher-order thinking and real world problem solving rather than just routine use of facts and procedures. If teachers were to aim for authentic student performance, then they would create assignments or assessment tasks that called upon students to construct their own meaning or knowledge, through in-depth disciplined inquiry, which addressed real world problems that had meaning beyond success in school.

In this study, nine standards were used to evaluate the quality of the teachers' assignments or assessment tasks: *depth of knowledge, knowledge criticism, knowledge manipulation, sustained writing, clarity and organization, connections to the real world beyond the classroom, supportive task framing, student control, and explicit performance standards/marking criteria.* Likewise, six standards were used to judge the quality of student work: *depth of knowledge, knowledge criticism, knowledge manipulation, sustained writing, quality of student writing/answers, and connections to the real world beyond the classroom.*

Under *depth of knowledge,* we conceptualized three types of knowledge: factual knowledge, procedural knowledge, and advanced concepts based on revised Bloom's knowledge taxonomy (Anderson & Krathwohl, 2001). It is important that we look at the

extent to which teachers require students to perform each type of knowledge in the day-to-day classroom assignments or assessment tasks. Higher-order thinking is captured by two standards, namely *knowledge criticism* and *knowledge manipulation*. *Knowledge criticism* is exemplified by tasks that ask students to compare and contrast different sources of information and to critique knowledge whereas *knowledge manipulation* is exemplified by tasks that demand students to organize, analyze, interpret, synthesize, and evaluate information; to apply knowledge and skills; and to construct new meaning or knowledge. In line with Newmann et al.'s authentic intellectual framework, *sustained writing* and *connections to the real world beyond the classroom* were also included. We believe that "what you test is what you get" and the abovementioned standards also apply to the evaluation of the quality of student work.

In particular, we contend that teacher's *supportive task framing* will result in higher intellectual quality in student work. Teacher's scaffolding of an assignment task, i.e., providing some structure and guidance, can assist students to accomplish a complex task (Nitko, 2004). *Task clarity and organization, student control, and explicit performance standards/marking criteria* are conceptualized based on Marzano's (1992) learning-centered instruction. The incorporation of these standards into the classroom assessment provides students with opportunities to engage in independent learning and critical thinking.

Methods

Samples

We collected 6,526 samples of teachers' assignments or assessment tasks and associated student work from the grades 5 and 9 lessons of English, Social Studies, Mathematics, and Science in 59 Singapore schools (30 elementary schools and 29 high schools) over two years (2004-2005). Grades 5 and 9 were chosen because students were streamed into different ability groups. In grade 5, there are three streams: EM1[1], EM2[2], and EM3[3] whereas grade 9 has four streams[4]: Special, Express, Normal Academic, and Normal Technical.

The types of assignments included daily class work, homework assignments, major assignments/projects, and teacher-made tests. Each teacher was asked to submit four samples each of high-quality, medium-quality, and low-quality student work in response to each type of assignment.

[1] EM1 students do well in English, one of the mother tongue languages (Chinese, Malay, or Tamil Language) and Mathematics

[2] EM2 students are slightly weaker in the mother tongue language

[3] EM3 students are weak in all three subject areas

[4] Students in the special and express streams have higher academic ability and will sit for their high-stakes national exam one year earlier than students in the normal academic stream. Students in the normal technical stream have lower academic ability and will sit for national exam that focuses on technical and vocational skills.

Scoring Rubrics

We developed two sets of generic scoring rubrics (one for assignments and the other for student work) and 20 subject-specific exemplars for both grade levels according to the authentic intellectual quality standards. All standards were scored on 4-point scales (ranging from 1 = no requirement/no demonstration to 4 = high requirement/high level). Thirty-five experienced teachers from non-participating schools were trained to be fully conversant with the standards and scoring rubrics. They were asked to try out their scoring on the subject-specific anchor papers for both assignments and student work. After their scoring had reached an interrater reliability of above .70, they were asked to score the actual assignments and student work samples. Throughout the scoring sessions, interrater reliability checks were conducted to ensure the integrity and consistency of scoring. Both teacher assignments and student work were randomly assigned to teacher scorers, and each standard was first scored independently by at least two scorers. The scorers compared their scores, and if they differed, they were asked to justify their individual scores and to discuss the discrepancy until they reached agreement on a final score. In scoring student work, the large volume of grade 5 English and Mathematics artifacts precluded double scoring for all of them. All the teacher assignments and student work in the other subjects were double-scored.

For both assignments and student work, scores for grades 5 and 9 were assigned on the basis of reasonable expectations within the grade level according to the subject-specific syllabus. In most of the subject areas, the percentages of exact agreement were above 70%, indicating good interrater reliability. Using the Many-Facet Rasch Modeling, statistical calibration was conducted to adjust rater severity. In all subject areas and grade levels, the reliability of separation index was low and the chi-square values were not statistically significant. This indicates that the overall differences between the raters (teachers) were not significant. The raters were consistent in their scoring and rater severity was relatively homogenous.

Results

Patterns of Classroom Assessment Practices

Over the two years, most of the teachers' assignments/assessment tasks in English, Social Studies, Mathematics, and Science at both grade levels were centered upon typical class work (see Table 1). Teachers who taught secondary Mathematics and Social Studies and primary Science had given more homework assignments compared to teachers who taught other subjects. Teachers' assignments comprised a relatively low proportion of major assignments or projects. In general, teachers gave more summative tests than formative tests.

Table 1. *Number of Artifacts across Subject Areas, Grade Levels, and Types of Student Work*

	Type of student work							
	Class work		Homework		Major Assignment/ Project		Test	
Subject area	Grade 5	Grade 9	Grade 5	Grade 9	Grade 5	Grade 9	Grade 5	Grade 9
English	888	356	97	66	11	-	-	54
Social Studies	351	127	54	76	-	5	-	12
Mathematics	522	255	55	167	-	-	60	81
Science	328	264	133	39	7	3	49	37

Quality of Teachers' Assignment Tasks

Table 2 shows the mean score differences of the quality of teachers' assignment or assessment tasks by grade level on each assessment standard. In general, the small effect sizes indicate that there were no practically significant mean score differences between the two grades. Although the assignment tasks in grade 9 required a slightly higher level of advanced concepts and more sustained writing, the nature of the tasks (i.e., a lot of worksheets and workbook/textbook exercises) still demanded that students presented knowledge as truth and reproduced what they had learnt from the textbooks. Teachers' task framing had focused more on procedural scaffolding so that students could arrive at the correct answers by using routine steps or procedures. Marking criteria were shared more explicitly with students at the grade 9 level.

Table 2. *Mean Score Differences of the Quality of Teachers' Assignment Tasks by Grade Level*

	Grade 5 n = 210		Grade 9 n = 136				
Standard	Mean	SD	Mean	SD	F	p	ω^2
Depth of Knowledge:							
Factual Knowledge	3.36	.80	3.30	.82	.46	.498	.00
Procedural Knowledge	2.61	1.03	2.54	.95	.41	.523	.00
Advanced Concepts	1.69	.80	1.83	.93	2.37	.125	.00
Knowledge Criticism:							
Presentation of Knowledge as Given	3.10	.99	3.48	.81	13.47	.000	**.04**
Compare and Contrast Knowledge	1.73	.83	1.82	.88	.78	.377	.00

Critique of Knowledge	1.54	.81	1.35	.77	4.85	.028	.01
Knowledge Manipulation:							
Reproduction	2.94	1.03	3.26	.92	8.66	.003	.02
Organization, Interpretation, or Evaluation of Information	2.27	.83	2.35	.87	.86	.355	.00
Application/ Problem-Solving	1.94	.90	2.04	.90	1.15	.285	.00
Generation/ Construction of Knowledge New to Students	1.68	.91	1.50	.83	3.31	.070	.01
Sustained Writing	2.18	1.16	2.49	1.06	6.58	.011	.02
Connections to the Real World beyond the Classroom	1.92	1.02	1.52	.86	14.47	.000	**.04**
Supportive Task Framing:							
Structure of the Task	2.27	1.00	2.25	.91	.03	.875	.00
Content Scaffolding	2.88	1.03	2.53	1.02	9.47	.002	.02
Procedural Scaffolding	1.88	.95	2.06	.99	2.81	.095	.01
Strategy Scaffolding	1.10	.38	1.17	.46	2.00	.158	.00
Clarity and Organization	3.48	.83	3.43	.81	.22	.639	.00
Learner Support:							
Student Control	1.55	.73	1.60	.70	.41	.522	.00
Explicit Performance Standards/Marking Criteria	1.15	.42	1.43	.70	22.84	.00	**.06**

Note. SD = standard deviation; n = number of teachers' assignment tasks.

The results for the quality of teachers' assignments across the subject areas at grades 5 and 9 are presented in Tables 3 and 4, respectively. For all subject areas except grade 5 Social Studies, the authentic intellectual quality of the teachers' assignment tasks was low. This is evidenced by higher mean scores on the elements subsumed under basic and rote knowledge, and lower-order thinking. The elements were factual knowledge, procedural knowledge, presentation of knowledge as given/truth, and reproduction. Mean scores for these elements ranged from 2.69 to 3.85 on a 4-point scale. In contrast, the mean scores were generally low (between 1.00 and 2.00) in all the elements that represent high authentic intellectual quality. An opposite pattern was observed in grade 5 Social Studies in which teachers' assignment tasks scored higher on most of the standards that measure high authentic intellectual quality. The mean scores were higher for the following elements: advanced concepts; critique of knowledge; organization, interpretation, synthesis, or evaluation of information; generation/construction of knowledge new to students; sustained writing; making connections to the real world beyond the classroom; strategy scaffolding; and student control. Although the authentic intellectual quality of assignment tasks in grade 5 Science was not high, the mean scores were found to be higher for the following elements: compare and contrast knowledge; organization, interpretation, analysis, synthesis, or evaluation of information; application/problem-solving; and procedural scaffolding. In grade 9 English, teachers' assignment tasks scored higher on three knowledge manipulation elements, i.e., organization, interpretation, synthesis, or evaluation of information; application/problem-solving; and generation/construction of knowledge new to students. In all subject areas and grade levels, the mean scores for clarity and organization were uniformly high, indicating that teachers' assignment tasks and instructions were mostly clear and well-organized.

Table 3. *Mean Score Differences of the Quality of Grade 5 Teachers' Assignment Tasks by Subject Areas*

	English n = 74	Social Studies n = 36	Maths n = 55	Science n = 45			
Standard	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	F	P	ω^2
Depth of Knowledge:							
Factual Knowledge	3.36 (.63)	3.03 (1.08)	3.85 (.41)	3.02 (.87)	13.84	.000	**.15**
Procedural Knowledge	2.91 (.76)	2.61 (1.02)	2.69 (1.10)	2.04 (1.11)	7.33	.000	**.08**
Advanced Concepts	1.50 (.63)	2.22 (1.17)	1.49 (.57)	1.80 (.76)	8.94	.000	**.10**
Knowledge Criticism:							
Presentation of Knowledge as Given	3.15 (.75)	2.06 (1.26)	3.85 (.45)	2.96 (.77)	36.74	.000	**.34**

Compare and Contrast Knowledge	1.57 (.68)	1.89 (1.14)	1.35 (.58)	2.36 (.65)	16.86	.000	**.18**
Critique of Knowledge	1.47 (.62)	2.08 (1.25)	1.05 (.23)	1.80 (.73)	16.95	.000	**.19**
Knowledge Manipulation:							
Reproduction	3.14 (.75)	1.97 (1.21)	3.67 (.67)	2.49 (.82)	35.63	.000	**.33**
Organization, Interpretation, or Evaluation of Information	2.04 (.71)	2.53 (1.08)	2.09 (.78)	2.64 (.68)	7.54	.000	**.09**
Application/ Problem-Solving	1.54 (.71)	2.42 (1.23)	1.76 (.64)	2.42 (.75)	16.24	.000	**.18**
Generation/ Construction of Knowledge New to Students	1.36 (.59)	2.89 (1.14)	1.27 (.56)	1.71 (.63)	45.27	.000	**.39**
Sustained Writing	1.92 (1.13)	2.86 (1.15)	2.20 (1.11)	2.02 (1.08)	6.10	.001	**.07**
Connections to the Real World beyond the Classroom	1.64 (.82)	3.11 (.92)	1.33 (.55)	2.18 (.98)	39.68	.000	**.36**
Supportive Task Framing:							
Structure of the Task	1.84 (.64)	2.58 (1.10)	2.78 (1.08)	2.09 (.95)	12.98	.000	**.15**
Content Scaffolding	3.24 (.76)	2.11 (1.26)	2.93 (1.09)	2.82 (.81)	11.35	.000	**.13**
Procedural Scaffolding	2.11 (1.02)	1.69 (.89)	1.40 (.68)	2.24 (.91)	9.92	.000	**.11**
Strategy Scaffolding	1.14 (.38)	1.28 (.66)	1.04 (.19)	1.00 (.00)	4.69	.003	**.05**

Clarity and Organization	3.32 (.86)	3.31 (.92)	3.73 (.68)	3.56 (.79)	3.29	.022	**.03**
Learner Support:							
Student Control	1.43 (.55)	2.11 (.89)	1.29 (.60)	1.62 (.75)	11.94	.000	**.14**
Explicit Performance Standards/Marking Criteria	1.04 (.26)	1.14 (.49)	1.33 (.51)	1.11 (.38)	5.47	.001	**.03**

Note. SD = standard deviation; n = number of teachers' assignment tasks.

Table 4. *Mean Score Differences of the Quality of Grade 9 Teachers' Assignment Tasks by Subject Area*

	English n = 38	Social Studies n = 14	Maths n = 41	Sciences n = 43			
Standard	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	F	p	ω^2
Depth of Knowledge:							
Factual Knowledge	3.32 (.70)	3.21 (.89)	3.95 (.22)	2.70 (.80)	25.31	.000	**.35**
Procedural Knowledge	2.71 (.80)	2.57 (.94)	2.83 (1.07)	2.12 (.82)	4.93	.003	**.08**
Advanced Concepts	2.24 (1.03)	2.29 (1.27)	1.63 (.77)	1.51 (.67)	6.53	.000	**.11**
Knowledge Criticism:							
Presentation of Knowledge as Given	3.32 (1.09)	2.93 (.83)	3.93 (.35)	3.37 (.62)	8.31	.000	**.14**
Compare and Contrast Knowledge	1.97 (1.08)	2.43 (1.22)	1.59 (.59)	1.70 (.67)	4.14	.008	**.06**
Critique of Knowledge	1.55 (.89)	2.43 (1.16)	1.00 (.00)	1.14 (.47)	19.47	.000	**.29**
Knowledge Manipulation:							
Reproduction	2.71 (1.14)	2.86 (.77)	3.85 (.48)	3.30 (.71)	14.51	.000	**.23**

Organization, Interpretation, or Evaluation of Information	2.87 (.94)	2.71 (.83)	2.12 (.68)	2.00 (.72)	10.50	.000	**.17**
Application/ Problem-Solving	2.71 (.90)	2.36 (1.08)	1.54 (.60)	1.84 (.69)	17.05	.000	**.26**
Generation/ Construction of Knowledge New to Students	2.26 (1.00)	1.79 (.89)	1.10 (.30)	1.12 (.32)	28.61	.000	**.38**
Sustained Writing	2.68 (1.21)	3.21 (.98)	2.24 (.94)	2.33 (.94)	3.92	.010	**.06**
Connections to the Real World beyond the Classroom	2.16 (1.08)	1.36 (.63)	1.12 (.40)	1.40 (.73)	13.15	.000	**.21**
<u>Supportive Task Framing</u>:							
Structure of the Task	1.95 (.77)	2.50 (.86)	2.88 (.93)	1.84 (.65)	14.65	.000	**.23**
Content Scaffolding	2.95 (1.09)	2.71 (.91)	1.98 (.88)	2.63 (.90)	7.39	.000	**.12**
Procedural Scaffolding	2.32 (1.21)	2.00 (.68)	1.80 (.78)	2.09 (1.00)	1.83	.145	.02
Strategy Scaffolding	1.29 (.69)	1.43 (.51)	1.05 (.22)	1.09 (.29)	3.84	.011	**.06**
Clarity and Organization	3.16 (1.05)	2.86 (.86)	3.78 (.53)	3.53 (.59)	7.43	.000	**.12**
<u>Learner Support</u>:							
Student Control	1.92 (.75)	2.43 (.76)	1.32 (.47)	1.33 (.47)	18.97	.000	**.28**
Explicit Performance Standards/Marking Criteria	1.24 (.43)	1.57 (.85)	1.24 (.44)	1.74 (.90)	5.58	.001	**.09**

Note. SD = standard deviation; n = number of teachers' assignment tasks.

.

Teachers' assignment tasks across streams at both grade levels were centered upon the mastery of basis knowledge and skills. The mean scores for factual knowledge, procedural knowledge, presentation of knowledge as given/truth, and reproduction were high in all streams. There were no significant stream differences for the quality of the teachers' assignment tasks in grade 5, except on the following three standards: factual knowledge, procedural scaffolding, and connections to the real world beyond the classroom. Upper stream assignment tasks appeared to focus more on factual knowledge, whereas lower stream assignment tasks were more likely on procedural scaffolding and provided slightly greater opportunity on making connections to the real world problems. There were some stream differences for the quality of teachers' assignment tasks in grade 9. Mean scores for the authentic intellectual quality elements were relatively low in the assignment tasks designed for NT students. NT assignments had much less emphasis on providing students with opportunities to engage in the learning of advanced concepts and higher-order thinking skills.

Quality of Student Work

The mean score differences by grade level were smaller than the mean score differences by subject area. The effect sizes were marginally significant. For both grade levels, student work demonstrated higher levels mastery of factual and procedural knowledge but limited understanding of advanced concepts. When students moved up to grade 9, their work tended to show higher levels of presentation of knowledge as truth and reproduction of factual knowledge, indicating that the authentic intellectual quality of the student work was low (see Table 5).

Table 5. *Mean Score Differences of the Quality of Student Work by Grade Level*

	Grade 5 n = 2555		Grade 9 n = 1542				
Standard	Mean	SD	Mean	SD	F	p	ω^2
Depth of Knowledge:							
Factual Knowledge	3.23	.83	2.86	.83	183.09	.000	**.04**
Procedural Knowledge	2.85	1.01	2.35	1.01	228.14	.000	**.05**
Advanced Concepts	1.65	.79	1.38	.60	135.98	.000	**.03**
Knowledge Criticism:							
Presentation of Knowledge as Given	3.14	.97	3.56	.74	220.06	.000	**.05**
Compare and Contrast Knowledge	1.67	.80	1.55	.67	26.95	.000	.01
Critique of Knowledge	1.45	.70	1.16	.44	218.95	.000	**.05**

Knowledge Manipulation:							
Reproduction	3.00	.99	3.41	.79	192.95	.000	**.04**
Organization, Interpretation, or Evaluation of Information	2.26	.76	2.03	.70	91.23	.000	.02
Application/Problem-Solving	1.89	.82	1.74	.67	38.35	.000	.01
Generation/Construction of Knowledge New to Students	1.61	.87	1.31	.57	140.02	.000	**.03**
Sustained Writing	2.27	1.19	2.42	1.08	16.15	.000	.00
Quality of Student Writing/Answers	3.01	.82	2.66	.87	165.52	.000	**.04**
Connections to the Real World beyond the Classroom	1.70	.94	1.55	.92	25.80	.000	.01

Note. SD = standard deviation; n = number of student work.

Tables 6 and 7 present the mean scores of the standards used to evaluate the quality of student work across all subject areas at grades 5 and 9, respectively. As found in the quality of teachers' assignment tasks, the authentic intellectual quality of student work in all subject areas except grade 5 Social Studies was low. Higher mean scores were noted in the following elements: factual knowledge; procedural knowledge; presentation of knowledge as given/truth; and reproduction. The mean scores ranged from 2.71 to 3.89 on a 4-point scale. In contrast, mean scores for all the elements representing high-intellectual quality were relatively low, ranging from 1.00 to 1.94. Compared to other subjects, student work in grade 5 Social Studies scored higher on the following elements: advanced concepts; critique of knowledge; application/problem-solving; generation/construction of knowledge new to students; and connections to the real world beyond the classroom. For grade 5 Social Studies, mean scores on factual knowledge; procedural knowledge; presentation of knowledge as given/truth; and reproduction were relatively low. Similar to the quality of teachers' assignment tasks, the mean scores in grade 5 Science student work, were higher on the following elements: compare and contrast knowledge; organization, interpretation, analysis, synthesis, or evaluation of information; and application/problem-solving. As expected, student work in grade 9 English scored higher on three knowledge manipulation elements: organization, interpretation, synthesis, or evaluation of information; application/ problem-solving; and generation/construction of knowledge new to students.

Table 6. *Mean Score Differences of the Quality of Grade 5 Student Work by Subject Area*

Standard	English n = 996 Mean (SD)	Social Studies n =405 Mean (SD)	Maths n = 637 Mean (SD)	Science n = 517 Mean (SD)	F	p	ω^2
Depth of Knowledge:							
Factual Knowledge	3.21 (.75)	3.14 (1.02)	3.56 (.72)	2.91 (.80)	64.76	.000	**.07**
Procedural Knowledge	2.74 (.90)	3.24 (.90)	3.42 (.85)	2.03 (.86)	273.07	.000	**.24**
Advanced Concepts	1.67 (.72)	2.14 (1.16)	1.33 (.53)	1.63 (.62)	95.53	.000	**.10**
Knowledge Criticism:							
Presentation of Knowledge as Given	3.22 (.82)	2.32 (1.25)	3.89 (.46)	2.71 (.67)	372.73	.000	**.30**
Compare and Contrast Knowledge	1.58 (.72)	1.66 (.98)	1.32 (.50)	2.29 (.73)	179.92	.000	**.17**
Critique of Knowledge	1.43 (.59)	1.98 (1.05)	1.00 (.00)	1.66 (.62)	229.64	.000	**.21**
Knowledge Manipulation:							
Reproduction	3.17 (.81)	2.29 (1.27)	3.72 (.61)	2.32 (.57)	396.83	.000	**.32**
Organization, Interpretation, or Evaluation of Information	2.10 (.66)	2.48 (.94)	2.16 (.77)	2.52 (.63)	54.35	.000	**.06**
Application/ Problem-Solving	1.66 (.70)	2.34 (1.23)	1.75 (.54)	2.17 (.70)	103.98	.000	**.11**
Generation/ Construction of Knowledge New to Students	1.41 (.68)	2.57 (1.25)	1.16 (.45)	1.74 (.57)	337.21	.000	**.28**

Sustained Writing	2.17 (1.17)	2.82 (1.13)	2.29 (1.23)	2.02 (1.06)	40.67	.000	**.04**
Quality of Student Writing/Answers	2.79 (.77)	3.18 (.85)	3.23 (.77)	3.03 (.86)	47.50	.000	**.05**
Connections to the Real World beyond the Classroom	1.36 (.74)	2.89 (.99)	1.28 (.45)	1.94 (.82)	501.58	.000	**.37**

Note. SD = standard deviation; n = number of student work.

Table 7. *Mean Score Differences of the Quality of Grade 9 Student Work by Subject Area*

	English n = 476	Social Studies n = 220	Maths n = 503	Science n = 343			
Standard	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	F	p	ω^2
Depth of Knowledge:							
Factual Knowledge	2.62 (.65)	3.00 (.90)	3.37 (.77)	2.37 (.65)	152.53	.000	**.23**
Procedural Knowledge	2.47 (.68)	2.09 (.95)	2.85 (1.15)	1.63 (.71)	133.03	.000	**.20**
Advanced Concepts	1.48 (.59)	1.49 (.76)	1.36 (.59)	1.19 (.46)	19.40	.000	**.03**
Knowledge Criticism:							
Presentation of Knowledge as Given	3.33 (1.04)	3.38 (.65)	3.94 (.24)	3.47 (.56)	74.74	.000	**.13**
Compare and Contrast Knowledge	1.62 (.78)	1.70 (.82)	1.38 (.50)	1.57 (.55)	16.76	.000	**.03**
Critique of Knowledge	1.32 (.56)	1.35 (.65)	1.02 (.13)	1.03 (.18)	73.21	.000	**.12**
Knowledge Manipulation:							
Reproduction	2.91 (.98)	3.36 (.61)	3.88 (.41)	3.44 (.57)	160.87	.000	**.24**

Organization, Interpretation, or Evaluation of Information	2.37 (.63)	2.05 (.72)	1.82 (.66)	1.87 (.64)	67.07	.000	**.11**
Application/ Problem-Solving	2.19 (.61)	1.53 (.64)	1.49 (.55)	1.61 (.61)	136.13	.000	**.21**
Generation/Construction of Knowledge New to Students	1.94 (.64)	1.04 (.20)	1.00 (.00)	1.08 (.27)	602.05	.000	**.54**
Sustained Writing	2.72 (1.25)	2.85 (1.06)	2.23 (.85)	2.02 (.91)	48.98	.000	**.09**
Quality of Student Writing/Answers	2.47 (.71)	2.31 (.85)	3.09 (.84)	2.51 (.90)	72.30	.000	**.12**
Connections to the Real World beyond the Classroom	2.41 (1.10)	1.05 (.21)	1.03 (.18)	1.42 (.68)	373.63	.000	**.42**

Note. SD = standard deviation; n = number of student work.

For all streams, student work demonstrated neither high level of mastery of advanced concepts nor higher-order thinking skills. Similar to the quality of teachers' assignment tasks, the stream differences on most of the standards for the quality of student work were not significant. Compared to other streams, EM1 student work in grade 5 demonstrated slightly higher mean score on factual knowledge. At grade 9, NT student work demonstrated lowest authentic intellectual quality.

Table 8. Correlations between the Quality of the Teachers' Assignment Tasks and the Quality of Student Work in Grade 5

	Student Work Quality				
Task Quality	Depth of Knowledge	Knowledge Criticism	Knowledge Manipulation	Sustained Writing	Connections to the Real World
Depth of Knowledge	.26**	.09**	.21**	.46**	.11**
Knowledge Criticism	.07**	**.78****	.68**	.21	.52**
Knowledge Manipulation	.13**	.72**	**.80****	.39**	.55**
Sustained Writing	.17**	.29**	.41**	**.70****	.28**

Connections to the Real World	.09**	.50**	.30**	.16**	**.80****
Supportive Task Framing	.01	.15**	.01	.30**	-.00
Learner Support	.09**	.26**	.33**	.36**	.33**

** $p < .01$

Table 9. Correlations between the Quality of the Teachers' Assignment Tasks and the Quality of Student Work in Grade 9

	Student Work Quality				
Task Quality	Depth of Knowledge	Knowledge Criticism	Knowledge Manipulation	Sustained Writing	Connections to the Real World
Depth of Knowledge	**.52****	.41**	.41	.64**	.16**
Knowledge Criticism	.16**	**.65****	.40**	.31**	.17**
Knowledge Manipulation	.20**	.61**	**.62****	.59**	.46**
Sustained Writing	.26**	.41**	.42**	**.80****	.25**
Connections to the Real World	-.05	.09**	.31**	.26**	**.72****
Supportive Task Framing	.09**	.24**	.28**	.34**	.10**
Learner Support	.04	.42**	.30**	.40**	.03

** $p < .01$

Tables 8 and 9 present the intercorrelations between the domain scores of the quality of teachers' assignment tasks and student work in grades 5 and 9, respectively. Most pairs of the correlations were statistically significant and had moderate to large correlations, indicating that the quality of the teachers' assignment tasks was correlated to the quality of student work. For example, when teachers' assignment tasks had required a high level of knowledge criticism, student work was most likely to demonstrate high levels of knowledge criticism and manipulation. This is consistent with the message 'What You Test is What You Get'. It also indicates that classroom assessment quality could have an impact on student learning and performance.

As expected, the correlations between the overall quality of the teachers' assignment tasks and the overall quality of the student work were strong and statistically significant at both grade levels (elementary schools: $r = .72$, $p < .01$ and high schools: $r = .63$, $p < .01$). The quality of teachers' assessment tasks matters in improving the quality of instruction and learning. The findings are consistent with Newmann et al. (1996) in the USA and Luke et al. in Australia (2000).

Educational Importance

The findings of this study have indicated that the teaching of higher-order thinking skills is still lacking in Singapore schools. The majority of the students have done well in rote learning and memorization. Teachers' assessment practices are mostly in the format of drill and practice of basic knowledge and skills so that their students can perform well in the high-stakes national exams and international assessments. The findings also set the stage for intervention plans of redesigning classroom teaching and assessment methods as well as efforts to enhance Singaporean teachers' assessment literacy through pre-service and in-service teacher professional development in alternative assessments and innovative task designs. Looking at both assignments and student work is an important strategy for encouraging teachers to be more reflective about their own classroom assessment and instructional practices. It will also deepen their reflections on the quality of an assignment and its impact on the nature of student work. Such strategy is useful in both pre-service and in-service teacher professional development.

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.

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.

Clare, L., & Aschbacher, P.R. (2001). Exploring the technical quality of using assignments and student work as indicators of classroom practice. *Educational Assessment, 7*(1), 39-59.

Darling-Hammond, L., & Falk, B. (1997). Using standards and assessments to support student learning. *Phi Delta Kappan, 79*, 190-199.

Frederiksen, J.R., & Collins, A. (1989). A district's approach to educational testing. *Educational Researcher, 18*(9), 27-42.

Lingard, B., & Ladwig, J. (2001). *The Queensland School Reform Longitudinal Study: Teachers' Summary*. Brisbane: Education Queensland.

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.

Matsumura, L.C. (2003). Teachers' assignments and student work: Opening a window on classroom practice. *CSE Report 602*. UCLA, Los Angeles: CRESST.

National Council on Educational Standards and Testing. (1992). *Raising standards for American education*. Washington, DC: Author.

Newmann, F.M., & Associates. (1996). *Authentic achievement: Restructuring schools for intellectual quality*. San Francisco: Josey Bass.

Newmann, F.M., Bryk, A.S., & Nagaoka, J.K. (2001). *Authentic intellectual work and standardized tests: Conflict or coexistent?* Chicago: Consortium on Chicago School Research.

Newmann, F.M., Lopez, A.S., & Bryk, A.S. (1998). *The quality of intellectual work in Chicago schools: A baseline report*. Chicago: Consortium on Chicago School Research.

Nitko, A.J. (2004). *Educational assessment of students*. 4th Ed. Upper Saddle River, NJ: Pearson/Merrill Prentice Hall.

Resnick, L. B. (1987). Learning in school and out. *Educational Researcher, 16*, 13-20.

Shepard, L.A. (1989). Why we need better assessments. *Educational Leadership, 46*(7), 4-9.