Assessment of 21st Century Skills in Singapore – A Decade's Journey

Poh Hua, TAY (tay_poh_hua@seab.gov.sg)

Siow Chin, NG (ng_siow_chin@seab.gov.sg)

Tian Lu, WEE (wee_tian_lu@seab.gov.sg)

Singapore Examinations and Assessment Board, Singapore

ABSTRACT

In 2004, the Singapore Ministry of Education (MOE) implemented Project Work at Grade 11 where students carry out an interdisciplinary group project to develop their thinking, communication and collaborative skills. This paper presents a methodology developed to describe and compare cohort acquisition of the skills assessed in Project Work over the past 10 years. Using a reference group, we developed a performance categorisation to classify students into meaningful and differentiable groups. A typical profile for each performance category is then fleshed out in terms of the criterion scores based on the assessment rubrics then augmented with a qualitative illustration derived from national standardisation exemplars. This methodology allows a comparison of cohorts' performance over the past decade, augmented with both quantitative and qualitative descriptions and provides an indication of students' acquisition of relevant skills articulated in MOE's 21st Century Competencies Framework 2010.

Sub-theme: Alignment between curriculum, instruction and assessment

INTRODUCTION

To help our students succeed in a fast-changing world brought about by globalisation and technological advancements, the Singapore Ministry of Education (MOE) has implemented Project Work (PW) at Grade 11 since 2004. This examinable subject required students to carry out an interdisciplinary group project over an extended period to develop their thinking, communication and collaborative skills (Bryer, 2006; Chong & Leong, 2014). PW entails students to draw knowledge and skills from across different subject areas to reinforce their understanding that cannot be compartmentalised, or seen in isolation. Students have to brainstorm and propose their own topic, plan and review their timelines, allocate and monitor areas of work, interact and collaborate with teammates of different profiles, and gather and evaluate primary and secondary research material. Using scoring rubrics provided by the Singapore Examinations and Assessment Board (SEAB), teachers assessed students based on specific assessment criteria that delineate the requirement to attain 'Approaching', 'Meeting' and 'Exceeding' Expectation for the following 3 components: (i) Written Report; (ii) Oral Presentation; and (iii) Group Project File. SEAB has also put in place a number of measures to ensure consistency in scoring and this includes giving all schools exemplar materials that illustrate the expected marking standards, conducting training for assessors and internal moderators and having internal and external moderation processes (Darling-Hammond, L., Herman, J., Pellegrino, J., et al., 2013).

In 2010, MOE produced a framework articulating the 21st Century Competencies (21CC) necessary for the globalised world. The three domains identified are: Civic Literacy, Global Awareness and Cross-Cultural Skills; Critical and Inventive Thinking; and Communication, Collaboration and Information Skills (MOE's *Standards and*

Benchmarks for Emerging 21CC; Jan 2014). Chong and Leong (2014) pointed out that the PW learning outcomes and model of assessment are closely related to the skills required of the workplace for the 21st century. These skills include creativity, critical thinking, problem solving, collaboration, independent learning and communication. In view of the strong association between 21st Century Competencies and PW outcomes, a methodology was developed to mine the PW data to infer students' acquisition of the relevant competencies as described in MOE's framework.

APPROACH

This paper presents a methodology to describe and compare Singapore Grade 11 cohort acquisition of the 21st Century Competencies assessed in PW over the past 10 years. The methodology comprises 5 key steps, abbreviated as 'D.E.P.T.H.':

1. **D**efine performance categories;
2. **E**xtract typical profiles quantitatively using assessment criterion scores;
3. **P**ortray typical profiles qualitatively using national standardisation exemplars;
4. **T**ie assessment criteria to relevant 21CC's domains and components; and
5. **H**ighlight performance trends of cohorts.

Step 1: Define performance categories. The first step is to identify a recent cohort to serve as reference group that facilitates the definition of performance categories. The reference cohort together with the PW Instrument is used to develop and define performance categories. Two guiding principles are adopted in the construction of the performance categories:

(i) *Substantive*: PW scoring rubric which is inherently criterion referenced provides indication of meaningful pre-existing categories that could be adapted.

(ii) *Statistical*: performance categorisation is constructed such that there are observable and meaningful differences between students in different performance categories. This can be achieved by maintaining at least a medium effect size difference between any two consecutive performance categories.

Based on the range of scores attainable by students within the Approaching, Meeting and Exceeding Expectation levels as stipulated in the Oral Presentation mark scheme (see Figure 1), we could identify some preliminary categorisation by computing the plausible scores across all the assessment criteria. The pre-existing categorisation include scores ranging (i) from 4a to 4b; (ii) from 4c to 4d; and (iii) from 4e to 4f.

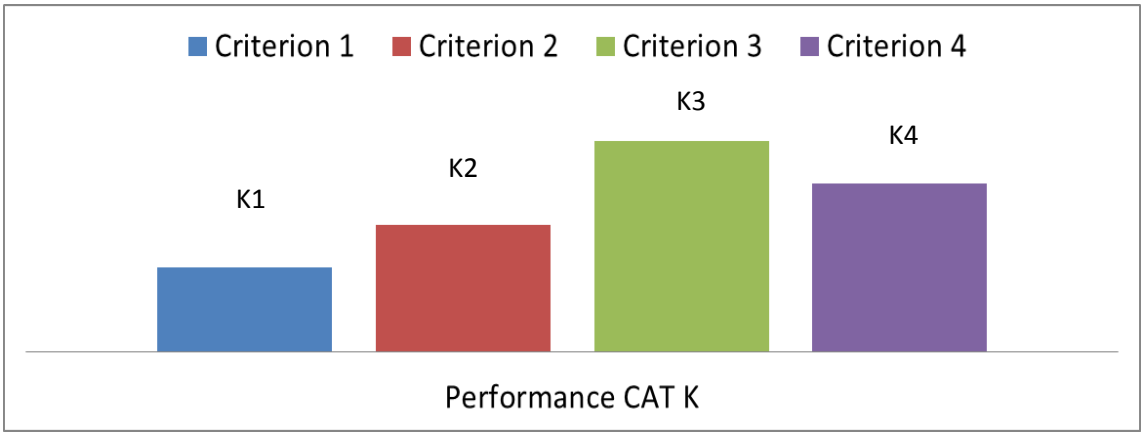
Figure 1

Assessment Criterion	Marks Allocation		
	Approaching Expectation (Individual Criterion mark range: a to b)	Meeting Expectation (Individual Criterion mark range: c to d)	Exceeding Expectation (Individual Criterion mark range: e to f)
Fluency and Clarity of Speech	Speaks haltingly and/or mumbles, and is difficult to understand at times	Speaks clearly and intelligibly most of the time	Speaks clearly and fluently throughout, at an appropriate pace
Awareness of Audience	Shows little awareness of audience	Shows some awareness of audience	Shows personal engagement with audience
Response to Questions	Answers are limited with little or no elaboration	Answers are relevant and contain some elaboration of ideas	Answers are relevant, well thought out and elaborated on
Effectiveness of Group Presentation	Presentation has limited effect due to lack of cohesion and organisation. Presentation aids do not enhance the presentation	Presentation is generally effective with some degree of cohesion and organisation. Presentation aids used appropriately to enhance presentation	Presentation is highly effective, cohesive and well-organised. Presentation aids used to effectively enhance overall effect
Plausible Score	Lowest: 4a Highest: 4b	Lowest: 4c Highest: 4d	Lowest: 4e Highest: 4f

Together with the statistical consideration that ensures at least a medium effect size difference between any two consecutive performance categories, we could classify students into a number of meaningful and differentiable performance categories: CAT1, CAT2, CAT3, … and CAT*N,* say.

Step 2: Extract typical profiles quantitatively using assessment criterion scores. In this step, we want to describe a typical student within each performance categorisation quantitatively in terms of the PW assessment criteria. In our study, we use the most probable student's assessment criterion scores within each performance category to serve as our typical profile. Figure 2 presents an illustration of the typical profile for performance category *K*, say. This profile suggests that students in this performance category are likely to attain a score of *K1* for assessment criterion 1 (Fluency & Clarity of Speech), a score of *K2* for assessment criterion 2 (Awareness of Audience) and so on.

Figure 2



Criterion 1 – Fluency and Clarity of Speech
Criterion 2 – Awareness of Audience
Criterion 3 – Response to Questions
Criterion 4 – Effectiveness of Group Presentation

Step 3: Portray typical profiles qualitatively using national standardisation exemplars. To further unpack the typical profiles established in step 2, each typical profile is augmented with a qualitative illustration derived from the national standardisation exemplars. The national standardisation exemplars are previously developed to complement the PW instrument with clear and descriptive illustrations of how students are expected to demonstrate at different competency levels across all the assessment criteria. In this step, the assessment criterion levels of each typical profile are respectively matched to the corresponding competency levels in the national standardisation exemplars. This allows us to extract an illustration of what students know and are able to do for each typical profile.

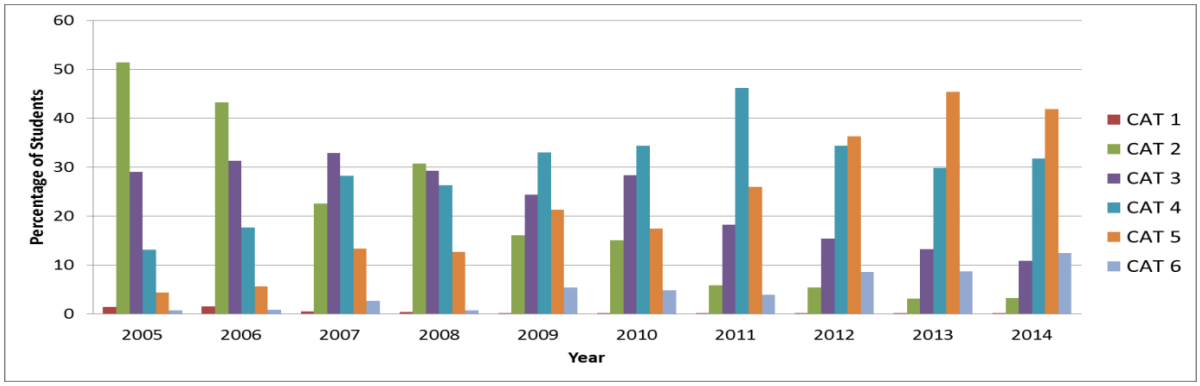
Step 4: Tie assessment criteria to relevant 21CC's domains and components. To provide linkages between PW and 21CC skills, the relevant domains and components from the MOE's Standards and Benchmarks document are first identified at the macro level. Next, appropriate assessment criteria from the different PW components are then linked to the relevant standards and benchmarks within the relevant 21CC domains and components.

Step 5: Highlight performance trends of cohorts. In the final step, distributions of the performance categories are used to compare the 10-year trend (from 2005 to 2014) of the PW cohort performance.

FINDINGS

As an illustration, this section demonstrates how the "D.E.P.T.H." methodology is applied to the Oral Presentation component in PW to ascertain the attainment of 21CC. Taking into consideration both substantive and statistical considerations mentioned in step 1, we found that students could be classified into six meaningful and differentiable performance categories. For instance, students in performance category 1 (CAT 1) are likely to have at least one criterion below meeting expectation while students in performance category 2 (CAT 2) are likely to achieve meeting expectation for all the criteria and etc. Students in each cohort are then systematically assigned to the 6 performance categories according to their Oral Presentation component scores. Figure 3 presents the distributions of the performance categories of the Oral Presentation component over the years.

Figure 3



We observed that the mode of each cohort has gradually shifted from CAT 2 in 2005 and 2006 to CAT 5 in 2012, 2013 and 2014. Moreover, the distribution of the performance categories of each year has shifted progressively to the right relative to the year before, indicating that the proportions of students attaining higher performance categories have increased over the years.

Comparing the competencies articulated in 21CC Standards and Benchmarks document and the PW assessment criteria, we found that the 'effective communication and collaboration' competency under the Communication, Collaboration and Information skills domain are relevant to the Oral Presentation component. These linkages enable us to infer the degree to which students are able to convey complex information and ideas coherently and clearly to influence and create impact for specific purposes and contexts. The 10-year trend of the Oral Presentation component indicates that oral presentation skills acquired by Grade 11 students over the years have progressed from CAT 2 to Cat 5.

With the qualitative descriptions of the typical profiles constructed in step 3, we are able to compare and contrast what students know and are able to do in the early years (2005 – 2007) and recent years (2012 – 2014). Segments of the typical profiles of CAT 2 and CAT 5 are briefly illustrated in the table below to show how students' oral presentation performances have evolved qualitatively from 2005 to 2014 for each assessment criterion (See Figure 1):

Fluency & Clarity of Speech	
CAT 2	**CAT 5**
Student's speech is mostly clear and intelligible, but pacing is uneven at times….	Student's speech is mostly clear, fluent and at a measured pace throughout the presentation….
Awareness of Audience	
CAT 2	**CAT 5**
Student shows some awareness of audience but is slightly delivery focused….	Student shows personal engagement with audience through a confident delivery, judicious use of smiles, sustained eye contact and effective hand gestures throughout the presentation….

CONCLUDING REMARKS

This study surfaced potential areas for review and enhancement of the PW instrument in addition to the information on the students' acquisition of the 21CC. For instance, in the process of linking PW's assessment criteria with the relevant domains and components from the 21CC Standards and Benchmarks document, we noticed that higher order assessment criteria could be incorporated into the current

PW rubrics to enhance the linkages between the PW instrument and the Standards and Benchmarks document. This is definitely useful information for future revision of the PW scoring rubrics.

In Singapore's national assessments, different facets of 21CC are being assessed across different subjects. The 'D.E.P.T.H.' methodology described in this paper could be readily adapted to 'mine' and describe students' acquisition of the relevant 21CC assessed in these different subjects. Although this methodology only requires existing data, it has the potential to provide new information regarding students' acquisition of the relevant 21CC at the national level. Owing to the fact that this methodology keeps the instrument intact, the inference of students' acquisition of the relevant 21CC may not be comprehensive. In the next step, our team will adapt and apply the 'D.E.P.T.H.' methodology to other subjects so that a more thorough picture of Singapore students' attainment of the 21st Century Competencies could be uncovered at key Grade levels (i.e. Grade 6, Grade 10, & Grade 12) over the years.

References

Bryer, K. C. (2006). *Pre-university project work in Singapore: An alternative mode of assessment in Singapore*. Paper presented at the 32nd International Association for Educational Assessment Conference, Singapore.

Chong, K. K. K., & Leong, S. C. (2014). *Assessment of 21st century skills: The Singapore experience with group project at grade 11*. Paper presented at the 40th International Association for Educational Assessment Conference, Singapore.

Darling-Hammond, L., Herman, J., Pellegrino, J.W., Abedi, J., Aber, J. L., Baker, E., … Steel, C. M. (2013). *Criteria for high-quality assessment*. Stanford, CA: Stanford Center for Opportunity Policy in Education. Retrieved from: http://edpolicy.stanford.edu/publications/pubs/847.

Ministry of Education (2014). *Standards and Benchmarks for the Emerging 21st Century competencies*. Ministry of Education, Singapore: Author.

Acknowledgement

We would like to thank our Singapore Examinations and Assessment Board (SEAB) management team for their support in this study; and our SEAB Assessment Specialists Ms Chong Kwei Kuen Karen, Ms Goh Wei Ling Christine, and Ms Lim Ing Yee Rachelle for their professional inputs to the qualitative descriptions of the typical profiles.