Title	Longitudinal teamwork measurement in Singapore Polytechnic: Initial steps and findings
Author(s)	Pee Suat Hoon and Chue Kah Loong

Longitudinal teamwork measurement in Singapore Polytechnic – initial steps and findings

Pee Suat Hoon[a]**and Chue Kah Loong**[b]

[a]Singapore Polytechnic,Singapore; [b]National Institute of Education, Singapore

Abstract

Teamwork skill is recognised by many governments and international agencies as an essential life skill and critical competency. In Singapore, teamwork is considered as a critical core skill by Skills Future SG and is reflected in MOE framework for 21st Century Competencies and Life skills Framework. In Singapore Polytechnic (SP), teamwork skill is identified as one of our graduate attributes. To holistically foster teamwork skill development, SP embarked on a poly wide educational teamwork research project in AY20/21. In this initial phase, the focus is on the systematic measurement of teamwork skill while the long term goal is to obtain psychometrically reliable and valid teamwork measurements to facilitate the development of teamwork skill in SP. This paper presents initial steps and findings of this longitudinal teamwork project.

Objective

Teamwork or collaboration skill is highlighted in many skills frameworks and identified by numerous educational institutions as a critical graduate attribute. Many employers expect graduates to possess good teamwork skill while graduates themselves expect to be adequately equipped with teamwork skills upon completion of their studies. To foster teamwork skill and other graduate attributes, many education institutions established graduate attribute statements and introduced related pedagogies into their courses. Singapore Polytechnic (SP) has similarly included teamwork as one of our graduate attributes. As can be seen in Figure 1, teamwork and other graduate attributes (Competency & Versatility, Communication & Collaboration, Creativity, Innovation & Enterprise, Ethics & Responsible Citizenry, Self-directedness & Personal Effectiveness and Global Mindset) sit right at the heart of our SP education model.

**CONTACT** Pee Suat Hoon, peesh@sp.edu.sg

Figure 1

The Singapore Polytechnic education model
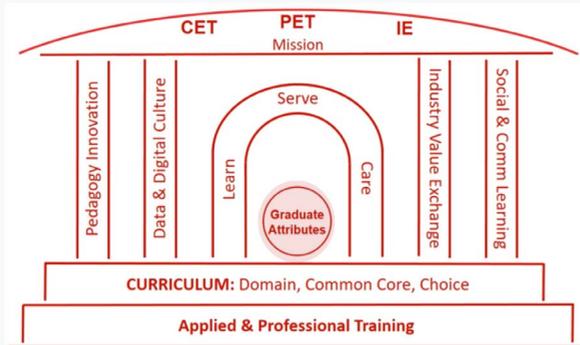
As teamwork skill cannot be learned in one module, it is usually developed through continuous practices with rich feedback and guidance. To effectively develop teamwork skill to the desired competency levels at SP, teamwork learning activities are carefully designed and integrated throughout the curriculum. There are three competency levels of teamwork graduate attribute at SP as illustrated by Figure 2, namely awareness, application and autonomy. At the awareness level, students are expected to raise their self-awareness of the teamwork attribute, begin acquiring knowledge and developing skills, abilities and behaviours under instruction, supervision and support. At application level, students continue to build, develop and refine their competency in the teamwork attribute. Students receive reduced guidance and support while gaining more room for self-exploration, adaptation and exercising judgement. At the autonomy level, students develop mastery and autonomy of the teamwork attribute. They are expected to demonstrate their awareness and application competencies with minimal or no guidance.

Figure 2

The competency levels of Singapore Polytechnic graduate attributes
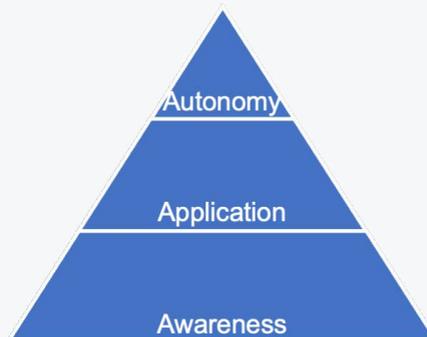
Curriculum and teamwork learning activities

The development of teamwork skill in SP is fostered by numerous collaborative exercises and teamwork projects in common core and domain modules. Zou & Ko (2012) showed that teamwork skills improved after Hong Kong University of Science and Technology students were systematically given explicit teamwork instructions, opportunities to practise and formative feedback across their three-year curriculum. Other researchers echo the usefulness of multiple and repeated teamwork learning experiences over the curriculum rather than within single modules (Siciliano, 1999).

The development of teamwork graduate attributes is supported by the SP curriculum which comprises 3 components: Domain, Common Core and Choice. Common Core curriculum is focused on developing desired skillsets and graduate attributes. There are 10 modules in the common core curriculum and three of these modules have included teamwork learning activities. They are Collaboration in Digital Age, Data Fluency and Sustainable Innovation Project. The Collaboration in Digital Age module provides a common starting point for all students to explicitly learn about teamwork and collaboration knowledge. Then, students undergo guided teamwork practices in the Data Fluency module through team based learning pedagogy and utilise teamwork skill in the context of Sustainable Innovation Project. In addition to the Common Core curriculum, teamwork skill has also been infused into numerous modules in Domain curriculum from earlier initiatives such as Conceive-Design-Implement-Operate (CDIO) and institutional graduate attribute project (Cheah & Sale, 2010).

While curriculum development is conducted at course level, assessment in SP is presently carried out at module level. Similarly, all teamwork assessments are embedded within respective modules in SP. This poses a challenge for the academic administration system to track the teamwork skill development across different modules to produce a course level aggregated teamwork skill achievement. To eventually assess and guide teamwork skill development across the curriculum, it is also useful to consider using a common standardised teamwork rubric so as to promote a common language across the curriculum.

Collated teamwork measurements across years of studies are very valuable to ascertain the teamwork skill competency attainment of students. However, these measurements, if of questionable reliability and validity, will not be useful. Teamwork measurements that are not reliable and valid cannot be used for decision making (Soland, Hamilton, & Stecher, 2013). Therefore, it is necessary to determine the validity and reliability of these teamwork measurements before they can be used in a summative way or for accountability purpose. This paper thus describes the first year of implementation and the focus is on the systematic measurement of teamwork skill.

Theoretical framework

Two frameworks are used in this study. First, planning and delivery, learning outcome and assessment are aligned at the module and course level (Figure 3). Related components are also aligned between the course and module levels. However, assessments are usually carried within the module and are not linked to the course. Hence, implementation of a course level longitudinal assessment will be shared in the remaining sections of this paper.

Figure 3

Alignment of planning and delivery, learning outcome and assessment at the module and course level
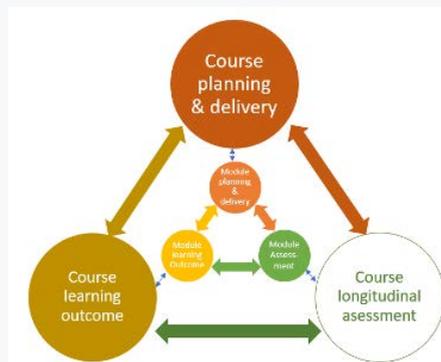
Second, the framework for teamwork measurements were adapted from the CATME instrument developed by Ohland et al. (2012). There are presently many teamwork instruments available to evaluate teamwork skills in education such as ITP Metrics (O'Neill et al., 2019), VALUE (Association of American Colleges and Universities, 2009), SPARKPLUS (Freeman & McKenzie, 2002), and WebPA (Willmot & Crawford, 2007). Among these tools, CATME developed by Ohland, et al., (2012) is most often used in educational institutions. It was found to have high predictive validity for peer ratings of teamwork (Enszer & Castellanos, 2013) and very high reliability values for the constructs (Loughry, Ohland, & Moore, 2007). In addition, it is reported that 42% of the students changed their behaviours in a positive way based on the feedback they received (Pung & Farris, 2011). This CATME measurement tool therefore has the added benefit of improving teamwork behaviours. Beyond its use in the academic community, Mayo (2016) has specially recommended CATME for validity studies in the corporate world.

There 5 categories measured in the CATME instrument are:
1. Contributing to the team's work
2. Interacting with teammates
3. Keeping the team on track
4. Expecting quality and
5. Having relevant knowledge, skills and abilities (KSAs)

Method

Schools were invited to participate in this teamwork project before the commencement of the academic year of 2020 (AY20/21).

Procedure
Measurement of students' self-assessment and peer assessment of teamwork skills was planned once a semester over their 3 years of studies. This worked out to six teamwork measurements for each student. Figure 4 shows an example of the teamwork data measurement plan for students enrolled in the Diploma in Civil Engineering (DCE). As the DCE course has not yet implemented the Common Core curriculum into their course in AY20/21, the data collection plan comprised predominantly Domain modules.

Figure 4.

Data collection plan for students in the Civil Engineering course
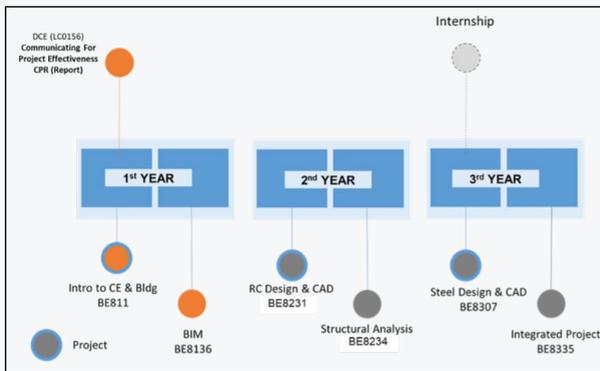
The project team from Civil Engineering course identified 6 modules with substantial teamwork learning activities across the 3 year course for the AY20/21 cohort of students. As illustrated in Figure 4, students carry out their teamwork assessments in BE8131 and BE8136 modules in Year 1, BE8231 and BE8234 in Year 2 and BE8307 and BE8335 in Year 3. In addition to domain modules from the respective diploma courses, the School of Life Skills & Communication (LSC) also contributed to the data collection efforts using their language modules. Conversations with the Department of Industry and Partnerships (INP) have started to include teamwork measurement data for the Internship module.

Measurement instrument

The categories in the CATME BARS (Behaviorally Anchored Rating System) instrument were used in this project instead of the CATME Likert-type instrument as there are 33 items in the short version and 87 items in the long version (Ohland et al., 2012). In using the rubric provided by the CATME BARS instrument, the number of questions was reduced to 5. This means that for a 4-person team, only 20 questions need to be answered with a BARS instrument as compared to 348 (= 87 x 4) questions for the long CATME Likert-type instrument. Scoring was done through peer assessment. Each category had a range of scores from 1 to 5, where a higher score indicated better teamwork skills. Figure 5 shows a sample of the descriptors for "Contributing to Team's Work" (adapted from Ohland et al., 2012).

In using the CATME survey questions throughout SP, all possible touch points across the curriculum use the same language and measurement tool. This helped to synergise

Figure 5

A sample of descriptors for "Contributing to the Team's Work"

Contributing to the Team's Work	5	5	5	5	5	• Does more or higher-quality work than expected. • Makes important contributions that improve the team's work. • Helps to complete the work of teammates who are having difficulty.
	4	4	4	4	4	Demonstrates behaviors described in both 3 and 5.
	3	3	3	3	3	• Completes a fair share of the team's work with acceptable quality. • Keeps commitments and completes assignments on time. • Fills in for teammates when it is easy or important.
	2	2	2	2	2	Demonstrates behaviors described in both 1 and 3.
	1	1	1	1	1	• Does not do a fair share of the team's work. Delivers sloppy or incomplete work. • Misses deadlines. Is late, unprepared, or absent for team meetings. • Does not assist teammates. Quits if the work becomes difficult.

the efforts across schools and enable students to develop their teamwork skills across six semesters. In alignment with SP's data and digital culture, a Learning Activity Management System (LAMS) was used to administer these CATME teamwork survey questions.

Sample

Due to the unforeseen COVID virus pandemic that broke out in Semester 1 AY20/21, most of the teamwork data measurements could not be conducted as students were unable to carry out teamwork learning activities that were originally planned in their lessons. In Semester 2, a total of 2051 empirical teamwork measurement data were made by Year 1 students from 9 courses as indicated in Table 1. Students conducted self-assessments (SA) and peer assessments (PA) of their team members.

Table 1

Year 1 students who carried out teamwork assessment in AY20/21

School	Courses	Sample size
Architecture & the Built Environment	Civil Engineering Integrated Events & Project Mgm Architecture	309
Chemical & Life Sciences	Chemical Engineering	55
Electrical & Electronic Engineering	Engineering with Business Aerospace Electronics	144
Business	Accountancy	591
Media, Arts & Design School	Media & Communication	228
Mechanical & Aeronautical Engineering	Mechanical Engineering	60
Life Skills & Communication		664
Total teamwork assessments completed		2051

Results

As the teamwork measurements gathered in Semester 1 was limited due to disruptions caused by COVID pandemic, Semester 2 measurements are analysed and discussed here.

Descriptive statistics

The teamwork assessment results for the 12 domain modules and 4 LSC modules for the five dimensions of teamwork assessment are provided in Table 2. Self and Peer assessment ratings are found to be quite comparable (average overall self-assessment = 4.47, average overall peer assessment = 4.50) with average peer rating minimally higher.

The lowest mean score for both self-assessment and peer assessment was "Keeping team on track". On the other hand, there was no observable pattern detected for highest mean score. The highest self-assessment rating was 4.50 (mean value) for two dimensions "Contributing to teamwork" and "Expecting quality" while the highest peer assessment was for the "Having relevant knowledge" dimension at 4.56 (mean value). The standard deviation obtained for both self and peer assessments was also found to be comparable.

Table 2

Descriptive statistics based on year 1 student teamwork measurements

Teamwork Dimensions	Mean Value	SD
SA Contributing to teamwork	4.50	0.61
SA Interacting with teammates	4.48	0.64
SA Keeping team on track	4.42	0.70
SA Expecting quality	4.50	0.64
SA Having relevant knowledge	4.48	0.64
PA Contributing to teamwork	4.55	0.59
PA Interacting with teammates	4.50	0.60
PA Keeping team on track	4.40	0.67
PA Expecting quality	4.49	0.62
PA Having relevant knowledge	4.56	0.57
Note: SA = Self assessment, PA = Peer assessment		

Figure 6 shows the same teamwork dataset using the Box and Whisker chart. It can be seen that the standard deviations for the self-assessment components are consistently wider than the peer assessments. This dataset offers a first glimpse of our students' teamwork performances across the various courses in different schools. However, it is premature to draw any final conclusions as further analysis will be conducted when more data are made available in the coming semesters.

While these initial results are useful, consideration must be given to the quality of these teamwork measurements. Quality teamwork measurement adds credibility and provides greater confidence when we use them. It is necessary to ensure that the teamwork measurements are of credible quality when we eventually award students their summative teamwork ratings. Similarly, these teamwork ratings need to be of certain quality if we need to measure the effectiveness of formative teamwork teaching and learning intervention.

Reliability

The two general criteria for evaluating psychometric measurements are reliability and validity (Jhangiani, Chiang, & Price, 2015). Reliability is a central concept and necessary condition for building a validity argument. Hence, it is useful to first establish the reliability of our teamwork measurements before proceeding to validity analysis. In this initial

Figure 6

Box and whisker plots of the 5 dimensions of teamwork measurements
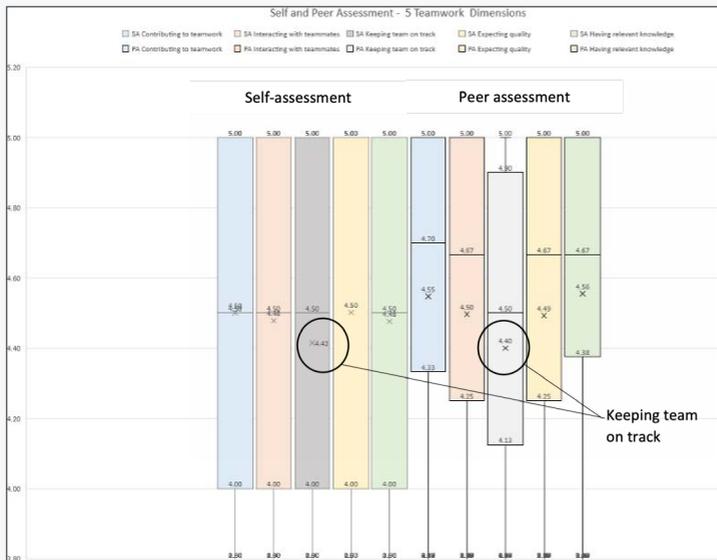
report, reliability analysis will be carried out and discussed in the following section.

Reliability refers to the extent to which measurements are consistent across other occasions of assessments, different editions of the test or different scorings by different raters (Livingston et al., 2018). In this research, reliability of teamwork assessments will be assessed in two ways; inter-rater reliability and test-retest reliability.

Inter-rater reliability in this research refers to the mean teamwork assessments for the same student made by different peers in two separate groups. This assesses the degree to which different groups of student raters agree in their assessment judgments of a particular student. Variations in peer assessment ratings may be attributed to the differences in subjective judgment of the peer raters, and may not necessarily imply that students had attempted to game the system. Furthermore, using the mean peer ratings for each group mitigates the subjectiveness of individual peers. Test-retest reliability examines how consistently students respond to their own assessment at different times. It is expected that an individual student would have the same perception of their own teamwork characteristics within a short time frame. For example, a student is likely to have the same level of Expecting Quality regardless of groupings and modules.

While inter-rater reliability is analysed using peer assessment ratings, test-retest reliability is analysed using the self-assessment of students evaluating themselves at different occasions within the same semester. Both inter-rater reliability and test-retest

Table 3

Inter-rater correlation using peer assessment

Dimension	Inter-rater reliability
1.Contributing to teamwork	0.417**
2.Interacting with teammates	0.434**
3.Keeping team on track	0.457**
4.Expecting quality	0.451**
5.Having relevant knowledge	0.429**
**p < .001	

reliability are computed by correlating between the two teamwork measurements. The correlation coefficients range from 0 indicating no consistency to 1 for perfect consistency between measurements. A correlation coefficient that is less than .3 is considered to be low while coefficient larger than .8 is considered to be high.

The sample data for computing the inter-rater reliability and test-retest reliability is taken from students who carried out teamwork assessments in two or more modules in semester 2. Inter rater reliability is computed using Pearson correlation coefficient as shown in Table 3. By correlating peer assessment from the two modules, all significant correlations are found to fall within the acceptable range of .417 to .457.

The test-retest reliability coefficient is obtained from the correlation of the students' self-assessment teamwork measurement made in 2 or 3 modules in the semester 2. By correlating self-assessment data from the two modules, it was found that all significant correlations fall in the acceptable range of .520 to .578 as shown in Table 4.

While significant correlations are obtained for inter-rater and test-retest in these teamwork measurements, more validity and reliability studies are planned and will be conducted on this existing set and new data.

Future directions

There were two notable observations from this initial study. First, the scores from the peer assessments were still relatively high, which could have been the consequence of several factors such as a misinterpretation of the descriptors or biasness in marking (e.g. friendship marking). In the event that a grade needs to be assigned to students as part of the teamwork graduate attribute, it would be difficult to discriminate between the various

Table 4

Test-retest correlation using self-assessment

Dimension	Inter-rater reliability
1.Contributing to teamwork	0.521**
2.Interacting with teammates	0.572**
3.Keeping team on track	0.578**
4.Expecting quality	0.520**
5.Having relevant knowledge	0.535**
6.Mean scores	0.641**

**p < .001

levels of teamwork. Second, although the content validity of the CATME instrument has been validated in previous studies (Ohland et al., 2012), further evidence of validity would need to be obtained as the instrument would be used on a large scale basis. As such, the next phase of the study would incorporate the use of a multi-facet Rasch model to resolve the two issues above.

Scholarly significance of the study

To the authors' knowledge, there are limited published examples of a longitudinal approach to teamwork skill development and measurement across the entire institution. While there are many teamwork research studies that are carried out for a module in a semester or year (e.g. Zou & Ko, 2012), it is relatively rare to find teamwork research of this scale in terms of time span and number of courses. Due to the difference in the scale of implementation, the design elements and issues that have to be taken into consideration are vastly different. This paper captured some initial steps and findings. More lessons will be shared in subsequent papers.

References

Association of American Colleges and Universities. (2009). *Inquiry and analysis VALUE rubric.* https://www.aacu.org/initiatives/value-initiative/value-rubrics/value-rubrics-inquiry-and-analysis

Cheah, S. M., & Sale, D. (2010, June). Sustaining Curriculum Innovation: The Diploma in Chemical Engineering Experience. In *6th International CDIO Conference* 14-18.

Enszer, J. A., & Castellanos, M. (2013, June). A Comparison of Peer Evaluation Methods in Capstone Design. In *2013 ASEE Annual Conference & Exposition* 23-32.

Freeman, M.A. and McKenzie, J., (2002) "Implementing and evaluating SPARK, a confidential web-based template for self and peer assessment of student teamwork: benefits of evaluating across different subjects", *British Journal of Educational Technology, 33* (5), 553-572.

Jhangiani, R. S., Chiang, I. A., & Price, P. C. (2015). Research methods in psychology-2nd Canadian edition: BC campus.

Livingston, S. A., Carlson, J., Bridgeman, B., Golub-Smith, M., & Stone, E. (2018). Test reliability-basic concepts. *Research Memorandum No. RM-18-01). Princeton, NJ: Educational Testing Service.*

Loughry, M. L., Ohland, M. W., & DeWayne Moore, D. (2007). Development of a theory-based assessment of team member effectiveness. *Educational and psychological measurement, 67*(3), 505-524.

Mayo, M. (2016). Workforce Assessments for the 21st Century: *Center for Curriculum Redesign.* https://curriculumredesign.org/wp-content/uploads/Workforce-Assessments-Report-March-2016-ARC-CCR.pdf

Ohland, M. W., Loughry, M. L., Woehr, D. J., Bullard, L. G., Felder, R. M., Finelli, C. J., Layton, R. A., Pomeranz, H.R, & Schmucker, D. G. (2012). The comprehensive assessment of team member effectiveness: Development of a behaviorally anchored rating scale for self-and peer evaluation. *Academy of Management Learning & Education, 11*(4), 609-630.

O'Neill, T., Larson, N., Smith, J., Donia, M., Deng, C., Rosehart, W., & Brennan, R. (2019). Introducing a scalable peer feedback system for learning teams. *Assessment & Evaluation in Higher Education, 44*(6), 848-862.

Pung, C. P., & Farris, J. (2011). A preliminary assessment of the CATME peer evaluation tool effectiveness. In *American Society for Engineering Education.* American Society for Engineering Education.

Siciliano, J. (1999). A template for managing teamwork in courses across the curriculum. *Journal of Education for Business, 74*(5), 261-264.

Soland, J., Hamilton, L. S., & Stecher, B. M. (2013). Measuring 21st century competencies. *Global Cities Education Network Report, 2013,* 1-68.

Willmot, P. & Crawford, A. (2007). Peer review of team marks using a web-based tool: an evaluation. Engineering Education: *Journal of the Higher Education Academy Engineering Subject Centre, 2* (1), 59-66.

Zou, T. X., & Ko, E. I. (2012). Teamwork development across the curriculum for chemical engineering students in Hong Kong: Processes, outcomes and lessons learned. *Education for Chemical Engineers, 7*(3), e105-e117.