International Journal of Engineering Education Vol. 36, No. 6, pp. 1850–1861, 2020
Printed in Great Britain

0949-149X/91 $3.00+0.00

Investigating the Reliability and Usefulness of Self- and Peer Assessments of a Capstone Design Project*

CHAN CHOONG FOONG[1] and PENG YEN LIEW[2,3]
[1] Medical Education and Research Development Unit (MERDU), Faculty of Medicine, University of Malaya, Malaysia. E-mail: foongchanchoong@um.edu.my
[2] Malaysia-Japan International Institute of Technology, Universiti Teknologi Malaysia, Jalan Sultan Yahya Petra, 54100 Kuala Lumpur, Malaysia. E-mail: pyliew@utm.my
[3] Process Systems Engineering Centre (PROSPECT), Research Institute for Sustainable Environment, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia. Email: pyliew@utm.my

There is increasing use of self- and peer assessments to assess behaviours of students working on group projects. This study aimed to explore the reliability and usefulness of self- and peer assessments during a capstone design project. A sample of 61 final-year undergraduate students aged 23 to 25 years old who were enrolled in Bachelor Degree of Chemical Process Engineering participated in the study. Students worked in groups of 5 to 6 members for 28 weeks to complete the project. Training was provided, and progress was monitored. Self- and peer assessments were conducted during the 6th, 14th and 22nd weeks. In each assessment, students rated their own behaviours and those of their peers using identical Likert scale questionnaires, and they also wrote feedback to themselves and their peers. Quantitative findings reported that, in the 6th week, students ranked themselves (mean = 3.98) significantly lower than how they ranked their peers (mean = 4.16). In the 14th week, students still ranked themselves (mean = 4.14) lower than how they ranked their peers (mean = 4.20). Last, in the 22nd week, students ranked themselves (mean = 4.24) equivalent to how they ranked their peers (mean = 4.24). For qualitative findings, feedback written to peers in the 22nd weeks was compared to self-assessment feedback from that week. Self- and peer observations on one's strengths and areas for improvement seem to converge both quantitatively and qualitatively towards the end of the project. It is also noted that both self- and peer assessment scores increased between the first and third assessments. The findings imply that students' behaviours improved while working on the capstone project. In conclusion, self- and peer assessments could be reliable and useful for chemical engineering students, and training students in how to conduct these assessments is essential to ensure successful implementation. Future qualitative research could identify how and why students gradually change their behaviours in long-term, team-based projects.

Keywords: peer assessment; self-assessment; capstone design project; chemical engineering; reliability; usefulness

1. Introduction

The Washington Accord adopts 12 graduate attributes to safeguard the standards of engineer education [1]. The graduate attributes explicitly describe the expected competencies in technical knowledge, technical skills and professional behaviours among graduates, and an engineering programme that meets the standard is accredited [1, 2]. The training in a four-year engineering programme, however, has compartmentalised this technical knowledge, these technical skills and these professional behaviours into courses (i.e., modules). Therefore, engineering design as a course is essential for students to utilise technical knowledge, technical skills and professional behaviours to solve industrial problems [3]. It is common for engineering programmes to implement a capstone design project for final-year or senior-year students. The capstone design project requires students to integrate all knowledge they have acquired from the engineering programme to solve a real-world industrial design problem. Capstone design projects are usually a form of project-based learning, in which students achieve a shared goal through collaboration in groups [4]. Students also take charge of their own learning; they outline the design and conduct investigations to achieve the project goals, and throughout the problem-solving processes, the students communicate and collaborate as they reflect and improve on their project [4, 5]. Apart from applying their technical knowledge and skills, students are required to demonstrate professional behaviours while working on the project. The Washington Accord defines these professional behaviours as ethics, teamwork and leadership, communication skills and lifelong learning [1]. In the real world, employers demand professional behaviours such as the abilities to communicate, to be a team player, to be proactive in solving problems and to show management skills [5–7].

A capstone design project requires group members to work collaboratively for months. This setting enables opportunities for development or examination of student behaviours. Traditionally, lecturers are the sole assessors of their students'

* Accepted 1 May 2020.

project-based learning. In the past decade, there has been increasing use of self- and peer assessments to assess behaviours of students working on group projects. This section discusses the validity, reliability and usefulness of peer assessments. First, validity refers to the extent to which the instrument can accurately measure what it intends to measure [8, 9]. Students must meet to propose, discuss, negotiate and make decisions on group projects. Through these interactions, peers have opportunities to observe the behaviours of their group members, and hence, peer observations could be more accurate than lecturers' assessments could [10, 11]. Second, reliability denotes the extent to which the instrument can consistently register the same measurements [8, 9]. Requesting group members to assess each other increases the number of assessors, and hence, this broader pool can enhance the reliability of the assessments [12]. The assessment is reliable if peers assign the same score to one student and if the score matches that on the student's own self-assessment. Third, usefulness indicates the extent to which an educational intervention could enhance students' graduate attributes [13, 14]. Some students are uncooperative when working in groups [15]. Through effective implementation of peer assessment, peer monitoring may prevent free riders and social loafing problems such as laziness among students [11, 16]. Peer assessments can also improve students' perceptions of team-based projects; in Willey and Gardner's study [17], students increased their engagement levels when working in groups and their abilities when working to achieve goals, and further, they practised higher-order thinking by observing and assessing their peers [18].

Self-assessment, in contrast, enables students to assess themselves, their actions and the consequences of their actions. Self-assessments engage students in the processes of observing themselves, producing judgements about themselves, and reacting to their judgements; these processes enhance self-efficacy [19]. Through self-assessment, students reflect on their strengths and areas for improvement [18]. By changing the roles from passive recipients of grades to active assessors, students can use assessments to monitor, guide and motivate them to achieve their goals [20]. Self-assessments enable students to be more responsible for their own learning [20] and are essential for students because self-reflection leads to continuing professional development independent of lecturers' instructions or input [21]. This is a form of lifelong learning. In the context of this study, a self-assessment is reliable if it registers the same score as that of the peer assessment.

Learning theories might explain how self- and peer assessments could be useful. Self-determination theory [22, 23] explains the development of professional behaviours among students using peer assessments. Students are extrinsically motivated when they obtain satisfaction by receiving rewards and avoiding punishments, when they are able to preserve their egos in front of the community and when they are able to recognise their importance to the group [22, 23]. When students are informed of conducting peer assessments for the capstone design project, as there are peer monitoring processes, students are extrinsically motivated to perform their tasks as team members. However, students may come to enjoy the group work experience and become intrinsically motivated to engage in teamwork. Students gain intrinsic motivation when they have pure interests (as compared to rewards and punishments) and enjoy their roles in the group [22, 23]. Reflection is like a mirror for students to view and focus on what they have behaved and respond emotionally to their behaviours [24]. Reflection consists of thinking processes in which individual students recall and evaluate their past actions to acquire a new understanding of their actions [25]. The students then confront their actions and resolve contradictions between their previous behaviours and what they could have done [24]. Triggered by these confrontations, students gain new understanding of their previous behaviours and adjust their behaviours accordingly [25]. Self-assessments are an application of reflection; they enable students to assess their actual behaviours against the expected behaviours and turn to a new leaf.

Peer assessments have benefits, but nevertheless, there are concerns about the validity, reliability and usefulness of both types of assessments. Students may be dishonest in assessing peers [26] or reluctant to criticise their peers [27], and this is a common concern especially among Asian students [28]. In self-assessments, students may be dishonest with themselves [26], and because self-assessment is a form of self-reflection, students must first have the ability to self-analyse [29]. Low- and high-performing students alike struggle to develop self-awareness [30]; most low-performing students are unaware that their performance is low, and they can overestimate their peers and themselves [20]. Lecturers must not assume that students have the capacity for self-reflection, and furthermore, overestimation and underestimation are both potential threats to self-assessments [10, 18]. The usefulness of self- and peer assessments is also of concern, as peer assessments may not be able to facilitate the positive development of a group member's behaviours due to other members' lack of authority [10]. A student may even become angry with other group members

upon receiving negative feedback. Peer assessments also assume students are be receptive to feedback given by peers, expecting them to be calm when receiving honest but critical feedback.

There are reviews that discuss the validity and reliability of self-assessments [19] and peer assessments [10], but there has been limited discussion on engineering students [7]. These previous studies may be unable to fully represent engineering students because engineering students have distinctive characteristics: they can be prone to make judgments objectively and impersonally rather than subjectively and personally, and they tend to work in a planned and decisive way rather than in a spontaneous and flexible way [31]. Students' characteristics influence their acceptance of educational training and assessments [32], and the question of how engineering students may respond to self- and peer assessments remains unexamined. The answer to this question is important because student acceptance of self- and peer assessments is essential to the effectiveness of these assessments [33, 34]. Therefore, an exploration of the reliability and usefulness of self- and peer assessments may guide future applications of both assessments for engineering students. The following sections feature this study's objectives, methodology and results; a discussion of the results; and a conclusion of the study.

2. Objectives of the Study

The objectives of the study were threefold: (i) to compare self-assessment and peer assessment results of final-year students in a capstone design project, (ii) to examine differences in self-assessment results of final-year students among three different phases in the capstone design project, and (iii) to examine differences in peer-assessment results of final-year students among three different phases in the capstone design project. The validity of the self- and peer assessments is discussed in the discussion section.

To the best of knowledge of the authors, this study is one of the pioneer studies on Asian students. An Asian perspective on the peer assessment may be different from the perspective of Western students due to cultural differences. Asian students do not resist group work but are more willing to compromise their own opinions rather than to argue with group members [35]. Criticising group members can be seen as a rude action that sabotages prestige of the group members. This cultural difference is consistent with the individualism–collectivism dimension; most Western countries uphold individualism, while most Asian countries favour collectivism [36]. In group project work, these two concepts manifest as follows: collectivism indicates student preferences such as 'We avoid conflicts' and 'We can do favours just to make our friends happy', whereas individualism indicates student preferences such as 'I will argue' and 'I do favours only for favours' [36]. Asian students tend to emphasise collectivism over individualism [37]; this means that students are in favour of avoiding conflicts to preserve the dignity of each group member and maintain harmonious relationships within the group. Asian students may suppress their actual thoughts due to the influence of collectivism [37], and consequently, peer assessments may be less reliable or useful for Asian students. Furthermore, Asian students may understand self-assessment differently: Asian students are taught to be humble, thus upholding the societal virtue of humility [38], and because it is rude for Asian students to rate themselves highly, low self-rating is a norm among these students [38]. In addition, because Asian students may be more timid than Western students are when representing themselves [39], self-assessments might be less reliable or useful for Asian students.

Cultural differences are an important aspect in learning enhancement. Cultural heritage influences the desires and approaches of students in self-improvement [40], and lecturers should consider the cultural backgrounds of their students to be able to design appropriate pedagogy [38] or to enhance the learning process [41]. This study's findings offer insight into the successful implementation of self- and peer assessments among Asian students.

3. Methodology

3.1 University, The Degree and Course Context

Sixty-one (n = 61) final-year (i.e., Year 4) undergraduate students who were enrolled in the Bachelor Degree of Chemical Process Engineering at the Malaysia-Japan International Institute of Technology (MJIIT), Universiti Teknologi Malaysia (UTM) Kuala Lumpur, participated in the study. These students were between 23 and 25 years old; 19 of them were male, and 42 were female. Students took the course SMJC 4343 (Chemical Process Design) during the September semester of the 2018/2019 academic year and then the course SMJC 4824 (Chemical Plant Design Project) during the subsequent February semester of the same academic year. The two courses were connected via a capstone design project, were conducted for two consecutive semesters (i.e., 28 weeks), exposed students to chemical process design principles and heuristics, and assigned students the task of designing a solution to a given real-world industrial problem. Besides technical knowl-

edge and skills, the courses aimed to enhance the professional behaviours of the students.

The capstone design project was divided into three phases (Fig. 1). During Project Phase 1, the student groups were required to produce a concept for the project, design process flow, select unit operations, formulate the operating conditions, calculate mass and energy balance, draw process flow diagram and design a waste treatment system. Meanwhile, the student groups continued the project in the subsequent semester for completion of the process control, process safety and profitability analysis (Project Phase 2). Next, the project entered the third phase covering individual equipment design (Project Phase 3), which was conducted as an individual assessment. In the two courses (i.e., capstone project), the same lecturer oversaw consistent student groupings who each worked on the same project. In brief, the Chemical Plant Design Project course was a continuation in every way of the Chemical Process Design course.

Students worked in groups of five or six members to produce the design of a chemical process plant. There were 12 groups, of which 11 groups consisted of 5 members and one group comprised 6 members. Each group featured students representing a diverse range of ethnicities, genders and academic performance levels; this mixture offered opportunities for the members to support one other, thus serving as training to establish mutual understanding in a real-world social setting [42]. Students from a variety of backgrounds working to achieve a shared goal is one of approaches to promote cooperative learning in classrooms [43].

Malaysia was ranked 44th among 56 countries worldwide in the individualism–collectivism dimension, a ranking that indicates the strong societal preference for collectivism [36]. Similarly, Malaysian people tend to 'save face' when interacting with their friends to maintain good relationships [44]. To continue the harmony of the student group, there was concern that students may give high ranks to peers to avoid conflicts among them. To prepare students for self- and peer assessment, the lecturer provided a briefing about the assessments in the early part of the semester, explaining the purposes of self- and peer assessments to students. As the capstone project was to be conducted across two semesters, the lecturer informed the students about the importance of imploring peers to monitor themselves. Students were expected to work outside the formal classes without the presence of the

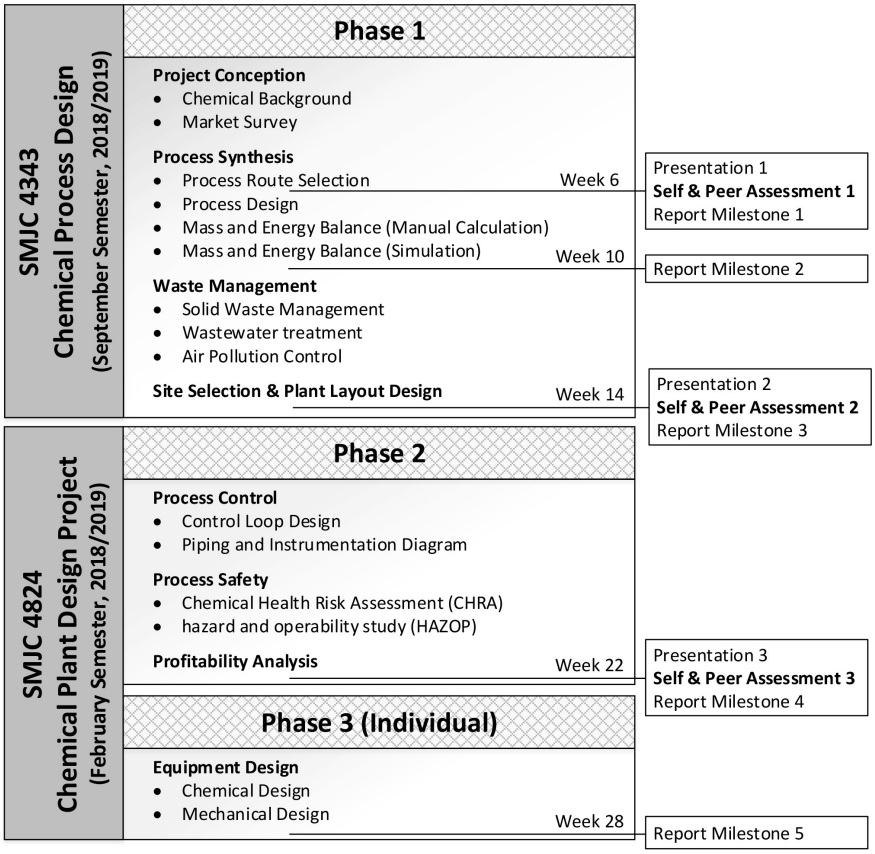

Fig. 1. Three phases of the capstone project.

lecturer and to practise self-reflection. Next, the lecturer provided training for the students on how to complete the self- and peer assessments using Google Forms.

Immediately prior to the administration of each assessment, the lecturer reminded students to be honest. Students were educated on the consequences of simply rating every group member highly, and the lecturer highlighted the importance of ethics to the students, both as adult students and future engineers. Students were reminded that peer assessments were to help peers develop appropriate behaviours for their future and not to simply make their peers feel happy by lying or stretching the truth. Students were also reminded that self-assessments were to help themselves to confront the truth about their behaviours and to make changes. This activity transcended good grades; students had to recognise that they were adults. The lecturer ensured that the students were well informed of the purposes and approaches to conduct project-based learning.

Cooperative learning is not just assigning students sit together and complete a task, group members must recognise that they are unable to complete the task unless every one cooperates [45]. Therefore, from time to time during the classes, the lecturer highlighted the behaviours of effective team members by telling short stories or sharing professional experiences about demonstrating appropriate workplace behaviours such as respect, cooperation, attentive listening and precision. This tactic was part of the hidden curriculum. Hidden curriculum refers to the unofficial or implicit values and behaviours which are conveyed in an educational setting [46]. The lecturer arranged in-class student discussions, walked from group to group, listened to each group and observed its dynamics, and asked questions (e.g., what was done) or sought clarification (e.g., who suggested what to be done). In addition to the lecturer who taught SMJC 4343 and SMJC 4824, there was one additional lecturer appointed to each group as the supervisor who helped to monitor project progress and group dynamics. Examples of monitoring included meeting with and intervening in specific groups that were observed to have conflicts, for instance, after reading their peer assessment results. Lecturers play an important role in managing student groups and promoting active and equal participation among the group members [4].

3.2 Self- and Peer Assessments

Self- and peer assessments were administered three times (i.e., in the 6th week, the 14th week and the 22nd week) as the capstone project was conducted. In addition to the self- and peer assessments, students as a group had to submit written reports or perform oral presentations. The setting offered opportunities for students to interact and work with group members, and this setting ensured that the self- and peer assessments were appropriately designed [10, 41].

During each series of assessments, students assessed themselves and all their peers in the same group. The questionnaire was a validated peer assessment instrument for use in measuring the professional behaviours of higher education students when working in groups [47]. In this study, the professional behaviours are confined to specific behaviours listed in Table 1. The instrument addresses many (but not all) professional behaviours, such as communication skills and teamwork, as defined by the Washington Accord. The version of Roberts et al.'s [47] instrument was modified due to three technical reasons. First, in all items, the 'tutorial' learning activities were replaced with 'discussion sessions' to accommodate the course context, and an example was the original item 1: 'this student prepared for tutorials' became 'this student prepared for the discussion sessions'. In the topic of discussion in the original item 8, 'case' was replaced with 'project', and the item became 'This student gave input that was focused and relevant to the project'. Second, specific examples were added in the original item 3 – 'this student showed behaviour and input that facilitated my learning' – to enhance clarity, and the item became 'this student showed behaviour and input (e.g., encouraging, responsive, cooperative) that facilitated my learning'. Last, the original item 10 was removed, as it was irrelevant to the purpose of the study ('During this block, working with my group facilitated my learning'). The modified instrument had nine items (Table 1), and each

Table 1. Instrument for peer-assessment

No	Statements
1	This student prepared for the discussion sessions
2	This student participated actively in the discussion sessions
3	This student showed behaviour and input (e.g., encouraging, responsive, cooperative) that facilitated my learning
4	This student was punctual to discussion sessions
5	This student listened to and showed respect for the opinion of others
6	This student brought new information to share with the team
7	This student was able to communicate ideas clearly
8	This student gave input that was focused and relevant to the project
9	This student accepted and responded to criticism gracefully

Table 2. Instrument for self-assessment

No	Statements
1	I prepared for the discussion sessions
2	I participated actively in the discussion sessions
3	I showed behaviour and input (e.g., encouraging, responsive, cooperative) that facilitate my learning
4	I was punctual to discussion sessions
5	I listened to and showed respect for the opinion of others
6	I brought new information to share with the team
7	I was able to communicate ideas clearly
8	I gave input that was focused and relevant to the project
9	I accepted and responded to criticism gracefully

item was a six-point Likert scale statement with the increments 0%, 20%, 40%, 60%, 80% and 100%, which denoted the time when there were discussion sessions. Defined criteria helped students objectively judge their peers [48].

In contrast, for self-assessments, all statements were identical (Table 2) except that the phrase 'this student' was changed to 'I'. Identical six-point Likert scales were applied for self-assessments.

In addition, for both self- and peer assessments, all students were asked to provide themselves feedback and provide their peers feedback using two open-ended questions. The questions were 'What attitudes/behaviours would you like this student to maintain in the future?' and 'What suggestions would you propose to improve the attitudes/behaviours of this student?' The written feedback represented self- and peer observation of one another.

3.3 Analysis

The data were analysed using IBM SPSS 23.0 software. Kolmogorov–Smirnov tests reported that the data were not normally distributed. Therefore, Wilcoxon signed-rank tests were used to compare the means of self- and peer assessment rankings of the same students. Next, Friedman tests and post hoc analyses were used to compare self-assessment scores measured at the 6th, 14th and 22nd weeks. The identical analyses were applied to compare repeated measures for peer assessment scores. Meanwhile, the gender influence on self- and peer assessments scores was deemed minimal in the study. Using Mann–Whitney U tests, female scores were not significantly greater than male scores for all self-assessments ($p > 0.01$) and all peer assessments ($p > 0.01$). This pre-analysis of demographic backgrounds was important to minimise threats to validity of the self- and peer assessment scores [20].

The qualitative analysis aimed to complement the quantitative findings. Student comments during the 22nd week were read and compared between what the students wrote to themselves and what they received from peers. Repeated vocabularies and keywords in the student comments were identified; relevant student quotes were collected and named as a behaviour. Student quotes were organised in a tabular form to present possible alignments between self- and peer observations.

4. Results

4.1 Quantitative Findings

Comparisons were made between self-assessment and peer assessment results of the students. In the 6th week, when the first self- and peer assessments were conducted, students ranked themselves (mean = 3.98) lower than what they received from peers (mean = 4.16). A Wilcoxon signed-rank test indicated that peer assessment scores were statistically significantly higher than self-assessment scores, $Z = -2.15$, $p < 0.05$. Next, in the 14th week, when second self- and peer assessments were conducted, students still ranked themselves (mean = 4.14) lower than their peers ranked them (mean = 4.20). However, a Wilcoxon signed-rank test indicated that peer assessment scores were not statistically significantly higher than self-assessment scores, $Z = -0.86$, $p > 0.05$. Last, in the 22nd week, when the third self- and peer assessments were conducted, students ranked themselves (mean = 4.24) equivalent to their peers' rankings (mean = 4.24). A Wilcoxon signed-rank test verified that peer assessment scores were not statistically significantly higher than self-assessment scores, $Z = -0.16$, $p > 0.05$. Statistical results are shown in Table 3.

Comparisons were made among the students' self-assessments in three phases. The means of self-assessment were 3.98, 4.14 and 4.24, respectively reported in the 6th, 14th and 22nd weeks during work on the group project. Comparison of the repeated self-assessment results was performed using Friedman's test showing a statistically significant increase, $\chi^2(2) = 12.29$, $p < 0.01$. Post hoc analysis with a Wilcoxon signed-rank test was conducted, and a significant increase was seen between the 6th week and the 14th week ($Z = -2.54$, $p < 0.05$), between the 14th week and the 22nd week ($Z = -2.43$, $p < 0.05$) and between the 6th week and the 22nd week ($Z = -3.21$, $p < 0.01$).

In contrast, the means of peer assessment were 4.16, 4.20 and 4.24, respectively reported in the 6th, 14th and 22nd weeks during work on the group project. Comparison of the repeated peer assessment results was performed using Friedman's test, showing a statistically significant increase: $\chi^2(2) = 8.85$, $p < 0.05$. However, post hoc analysis with a Wilcoxon signed-rank test was conducted, and the

Table 3. Results of the self- and peer-assessments

	Self-Assessment		**Peer-Assessment**	
	Mean ± SD	**Min; Max**	**Mean ± SD**	**Min; Max**
6th week	3.98 ± 0.60	2.67; 5.00	4.16 ± 0.39	2.81; 4.83
14th week	4.14 ± 0.60	2.67; 5.00	4.20 ± 0.41	2.83; 4.81
22nd week	4.24 ± 0.64	2.56; 5.00	4.24 ± 0.46	2.39; 4.86

Table 4. Possible alignments between self- and peer observations on what behaviours should be maintained

Student Name	**Behaviour to Maintain**	**Myself**	**Peer 1**	**Peer 2**	**Peer 3**	**Peer 4**
Ray	Hardworking	Maintain my diligence and quality of works	Work attitudes	Finished his works fast [. . .] hardworking	Hardworking [. . .] you will be the first person that finishes your own part	–
Lily	Hardworking	Hardworking	Hardworking	Her dedication	Hardworking	–
Sally	Hardworking	Hardworking	Prepared for meeting	Maintain your hardworking and work oriented attitudes	This student gave tremendous amount of efforts towards completion of the project	Can finish her works and helping others
Noon	Hardworking	You are hardworking	Hardworking	Very hardworking	Very hardworking	–
Urish	Hardworking	Hardworking	Hardworking	Her commitments	Submit works following the dateline	Good work quality
Hanks	Contribute positive energy to the group	The positive vibes and always smile	Positivity	Positive attitudes	Maintain his optimism because this is the trait that should exist in a group-based project	Always positive during hard times [. . .] always give encouragements
Fonny	Helpful	Helpful person and always show your efforts to support your teammates	Very helpful person	Always helps each other	Has the wills to help other group members	–
Amy	Share ideas/ information	Keep up the spirit of finding information in completing the works	Always come out with new ideas	Really good in giving opinions and suggestions	Always give ideas to solve problems	Being very helpful all the time
Linsey	Leadership	Thinker	The leadership and good in command	Leadership	The way you work hard to achieve something that gives benefits to the group, makes me want to keep working hard to improve and prepare myself in the future	You are a very good leader

increase was not significant between the 6th week and the 14th week ($Z = -1.36$, $p > 0.05$), between the 14th week and the 22nd week ($Z = -1.83$, $p > 0.05$) and between the 6th week and the 22nd week ($Z = -1.93$, $p > 0.05$).

4.2 Qualitative Findings

Table 4 presents possible alignments between self- and peer observations on what behaviours should be maintained in the future as of the 22nd week. These behaviours were hardworking, contribute positive energy to the group, helpful, share ideas/ information and leadership. Student quotes were inserted into the table as supporting evidence. Students were given pseudonyms.

Table 5 presents possible alignments between self- and peer observations on what behaviours should be changed in the future as of the 22nd week. The behaviours were shyness, confidence, participation in meetings, punctuality, teamwork and leadership. Student quotes were inserted into the table as supporting evidence.

Table 5. Possible alignments between self- and peer observations on what behaviours should be changed

Student Name	Behaviour to Change	Myself	Peer 1	Peer 2	Peer 3	Peer 4
Lily	Shyness	Communication skills	Be more brave and [speak] louder	Don't be too shy in giving your ideas	To be able to speak the ideas	Too quiet
Fonny	Confidence	Need more practices in terms of communication and presentation	Be confident in sharing new ideas	Be more confident on the results of your own creativity and ideas	Give more ideas during discussion	–
Sam	Confidence	Present ideas more clearly	Have confidence in yourself	Really grateful to be as your team mate but [please] be brave	Keep up the confidence levels	–
Sally	Participation in meetings	Be more active during meeting session	Communicate more with team members	Try to communicate more [with] team mates instead of doing task alone	Communicate better	Participate actively in the discussion, share your workload with others
Nick	Participation in meetings	Responsive in group meetings	Be more focused during discussions	Playful during meetings which should be avoided in the future	Respond to WhatsApp group	Let's be more active and share ideas clearly in discussions
Don	Participation in meetings	Focus more on the project	Active in discussions	To be more prepared during meetings	Be more committed in doing tasks	Be more prepared and pay attention on discussions
Alfredo	Punctuality	Come to meetings early	Attend meetings on time	Be more punctual, do not procrastinate, focus on the group project and please do your works on time	Be more punctual for the discussions	Be more punctual and do not delay your works until the end and make others waiting
Vinnie	Teamwork	Prepared for meeting	Be responsible towards your works and do not simply give up, and depend on others to settle your own work [. . .] do your work as a team	She delegated task to teammates and ended up disastrous for some of group members [. . .] she also focused more on her final year project than paying attention [. . .] I believe she can do more when it comes to group work	Do not escape from meetings [. . .] do not make others to do your tasks [. . .] do not burden your group members, please contribute	–
Michael	Leadership	I might not be the ideal leader	Always be someone who can accept people ideas	Be more approachable and try to accept others' points of view	Try to consider others' ideas	–
Ray	Leadership	Improve my trust on my teammates' works and to discuss with them before doing any amendment	Alert to the works of other people	We all know that you are hardworking but please tell us which parts that you did because [we] don't want [our] works to become redundant	Team work and the task delegation	Please do not just focus on your works without concerning the tasks and time of other people

5. Discussion

5.1 Reliability of Self- and Peer Assessments

There is a significant difference between the first self- and peer assessment scores. Threats to reliability seem to be true, lack of skill or dishonesty in the peer assessment is possible and overestimation and underestimation in the self-assessment may have occurred. However, these differences decrease as the students approach the end of the capstone project. By the 22nd week, there is no significant difference between self- and peer assessment scores. The comparison implies that, at the end of the capstone project, students' opinions of themselves were consistent with what their peers observed. The alignment between self-reflection feedback and

feedback from peers is evidence that positive changes in professional behaviours gradually occur during the project.

Prior to conducting this study, there was concern that Asian students might underestimate themselves due to humility and timidity. The concern appears to have been unwarranted, however, as this study demonstrated that students who are humble do not necessarily doubt their abilities [49]. Initially, the peer assessment scores significantly differed from those of the self-assessments, but there was a gradual convergence of these scores. The lecturer provided essential and successful training, and appropriate training helps students to assess themselves and peers accurately [48]. With appropriate training, students recognise that honest, descriptive and constructive feedback to peers is optimal [41], and for self-assessments to work effectively, students must recognise that honest self-reflection is crucial to self-improvement. Students need to be well informed that the purpose of peer assessments is not to make peers feel bad but rather to encourage peers to improve themselves. In addition, the training is not limited to students and that lecturers must incorporate training in their own practices. Lecturers also must be trained to develop their abilities in designing appropriate training for students [50]. As conflicts during group work are inevitable, lecturers must carefully resolve interpersonal problems among students [50].

Prior to conducting this study, there was also concern that Asian students, due to their preference of collectivism over individualism, might tend to give favourable comments to peers in order to avoid conflict [36]. However, the qualitative findings of this study suggest that the students were honest in indicating or recommending areas for improvement (e.g., 'be punctual', 'be confident', 'please contribute') to their peers. In comparison to studies on Western students, an Australian study reported that international, Asian engineering students were likely to rank peers more generously than local students were [28]. The researchers hypothesised that these students were reluctant to criticise peers due to culturally based discomfort. In the United States, without being formally trained, pharmacy students ranked themselves significantly lower on self-assessments than their peers ranked one another [51]. Another study in United States also showed that peer assessment results were significantly higher than self-assessment scores among dentistry students [52]. To conclude, it appears to be training, rather than Asian and Western cultural differences, that influenced the scoring disparities. It is possible that regular interventions by the lecturer to promote the hidden curriculum during classes may have convinced the students to provide honest feedback to themselves and peers.

Nevertheless, one shortfall of the training in this study was that some student feedback was non-constructive. Examples of vague and irrelevant feedback include single-word comments ('Good') and negative sentiments ('Her ignorance to occasionally respond to group's WhatsApp is what annoys me'). Therefore, future training should inculcate in students the need to provide effective feedback. Elements of effective feedback are descriptive (e.g., 'You did not complete the work on time') instead of judgemental (e.g., 'You are lazy') and constructive (e.g., 'You could set phone reminders') instead of complaintive (e.g., 'You are late') [53]. In addition, peer assessments may be made anonymous to create a non-threatening atmosphere for students to give honest feedback. Nevertheless, anonymity must not be misused.

After 22 weeks of collaboration, self-observation did not differ significantly from peer observation. The findings imply that students could have developed more accurate insight into their performance after engaging in repeated self- and peer assessments. One of the doubts about implementing self-assessments was whether students would have had insight to evaluate their performance [20, 30], but this concern is now resolved. Ensuring reliability of self-assessments (in comparison to peer observation) is essential because, to promote lifelong learning and self-reflection, students themselves must have accurate awareness of their own performance [30]. Insight is important for students to prevent academic failures and for graduates to recognise the need for continuing professional development as they progress in their careers [30, 54].

5.2 Promotion of Professional Behaviours

Both self- and peer assessment scores increased throughout the two semesters; students improved in their behaviours while working on the capstone project. In one example of feedback, a peer commented about a student named Zack that his 'Teamwork had improved compared to last semester'. These findings echo Willey and Gardner's study [17] in which self- and peer assessments could engage students to collaborate when working on group projects.

Interpreting Boud, Keogh and Walker's [25] model of reflection, self-assessments enable students to recall their experiences collaborating and communicating with group members during meetings. Next, positive (e.g., fruitful discussions) or negative events (e.g., conflicts) encourage students to attend to their feelings. In addressing these conflicts, students may feel upset at first but may

also ask themselves why they are treated as they are treated. Is it their own fault or that of their peers? In the first self-assessment (i.e., in the 6th week), Jane wrote, 'I was pretty intimidated by the project at the beginning so I am aware [that] I may have not come up with the best ideas for the project, but slowly [I have] got a grip. So, I would continue to fulfil all my tasks responsibly until the project is completed'. Students must re-evaluate their experiences to maximise the positive and to minimise the negative, and in doing so, they gain new perspectives of themselves as collaborative group members and become willing to change their behaviours [25]. It is never too late for self-reflection. During the third self-assessment (i.e., in the 22nd week), Michael told himself that 'Truthfully, I could be blamed on the lack of output [. . .] I might not be the ideal leader [. . .] I should reflect on the mistakes even the tiny one that I had done during this final year and bounce back from it.' Self-reflection is a driving factor in lifelong learning, which is a graduate attribute listed in the Washington Accord. This study provided an early intervention that illustrated the usefulness of self-assessment for self-reflection. Future studies could consider using structured guides for students to reflect on their self- and peer assessment results. Low-performing students usually need specific guidance to recognise actionable areas for improvement [54]. While theories of reflection are abstract, a three-step model (i.e., 'What do I feel about the feedback', What do I think about the feedback', and 'What actions could I take to improve') may be practical and easy for higher education students to use [21].

Self-determination theory [22] could explain the increasing scores among peer assessment results. Knowing that they will be assessed by peers, students are extrinsically regulated to show positive behaviours when working with peers on the capstone project. As members of a group, they may want to avoid feeling guilty or to exhibit ego or pride. Furthermore, based on the written feedback to themselves and to their peers, there are words of encouragement, enjoyment and satisfaction about group work. This phenomenon suggests that group members may have developed genuine friendship throughout the capstone project. Examples of this camaraderie include messages such as 'I love your team spirit', 'Don't cry alone [. . .] we [are] always with you' and 'Take care of health'. Once students internalise these previous external regulations into intrinsic motivation, they enjoy and gain satisfaction from being a part of the team [22].

In conclusion, self-assessment potentially promotes self-reflection among students. Though peer assessments may provide extrinsic motivation to initiate non-voluntary learning processes, intrinsic motivation is the ultimate goal. The two assessments could complement each other in developing professional behaviours among engineering students.

5.3 Limitations of the Study

This study has limitations. First, peer assessments are considered valid because peers have more opportunities than lecturers have to observe student behaviours over a 28-week, team-based project, but future studies could compare self- and peer observations with the results observed by the lecturers. This comparison could further verify the validity of self-, peer and lecturer observations. Second, qualitative investigation of the study was limited. Future studies should consider conducting focus group discussions with students to investigate how and why they have changed their behaviours over the project period. Some questions remain unanswered in this study: Were students angry or hurt when reading their peer assessments results? How did these students manage their negative emotions? Why did they change? In addition, the Asian perspective requires re-examination through prompting students to respond whether they honestly evaluated their peers and themselves and, if applicable, why they were not honest. For future research, interviewing the students in a safe environment and establishing good rapport with interviewees could prove integral to collecting honest responses [55]. Third, the professional behaviours of engineering students are broader than the relatively few options that this study assessed. For instance, element of lifelong learning was not measured. Because an appropriate instrument is yet to be developed for engineering education based on the Washington Accord, future studies could consider adding relevant behaviours for measurement or combining existing validated instruments. This investigation would provide a more holistic measurement of professional behaviours among engineering students.

Fourth, it was found that feedback written to peers and themselves were often less than 15 words in length; in this study, students received no training on how to write proper feedback in the study. Future interventions should teach students to be specific, enforce the positive, make feedback descriptive rather than evaluative and constructively give feedback about areas requiring improvement [21, 53]. Writing effective feedback for students themselves and peers are relatively soft skills in comparison to the essential knowledge for the field of chemical engineering. Therefore, interdisciplinary collaboration among engineering educators and education specialists would be practical with the aim to conduct appropriate training for students. Future interventions should train stu-

dents to provide detailed feedback to themselves and their peers. Detailed feedback could provide rich data for qualitative analysis. Last, despite the fact that this was a pioneering study in an Asian region, this study involved only one cohort of students in a single institution. Future research with a larger sample size involving multiple institutions would enhance the generalisability of findings.

6. Conclusion

In this study, self- and peer assessments were repeatedly conducted throughout a capstone project. The self- and peer assessments diverged at the beginning of the project. At the end of the project, self- and peer assessments were found to be reliable, and these assessments were also useful for engineering students to enhance their professional behaviours. Meanwhile, regular student training and monitoring were essential in implementing the assessments. The findings recommend a broader application of self- and peer assessments for engineering capstone design projects. Despite the distinctive characteristics of engineering students and students from Asian countries, the study did not find an observable difference in the effectiveness of these assessments in these two demographics. However, experience of this study was preliminary, and future endeavours should provide a greater understanding of the elements influencing student behaviours when working on group projects. This greater understanding would help educators train engineering students to meet the Washington Accord's expectations on their professional behaviours.

References

1. International Engineering Alliance Secretariat, *25 Years of Washington Accord: 1989–2014 Celebrating international engineering education standards and recognition*, International Engineering Alliance Secretariat, Wellington, New Zealand, 2014.
2. A. A. Aziz, M. J. M. M. Noor, A. A. A. Ali and M. S. Jaafar, A Malaysian outcome-based engineering education model, *International Journal of Engineering and Technology*, **2**(1), pp. 14–21, 2005.
3. C. L. Dym, A. M. Agogino, O. Eris, D. D. Frey and L. J. Leifer, Engineering design thinking, teaching, and learning, *Journal of Engineering Education*, **94**(1), pp. 103–120, 2005.
4. D. Kokotsaki, V. Menzies and A. Wiggins, Project-based learning: A review of the literature, *Improving schools*, **19**(3), pp. 267–277, 2016.
5. J. Uziak, A project-based learning approach in an engineering curriculum, *Global Journal of Engineering Education*, **18**(2), pp. 119–123, 2016.
6. L. Wilson, S. Ho and R. H. Brookes, Student perceptions of teamwork within assessment tasks in undergraduate science degrees, *Assessment & Evaluation in Higher Education,* **43**(5), pp. 786–799, 2018.
7. K. P. Nepal, A relative relevance approach to refine inconsistentpeer-and self-assessment scores in teamwork assessment, *International Journal of Engineering Education*, **34**(4), pp. 1289–1298, 2018.
8. S. S. Salmond, Evaluating the reliability and validity of measurement instruments, *Orthopaedic Nursing*, **27**(1), pp. 28–30, 2008.
9. V. A. Scholtes, C. B. Terwee and R. W. Poolman, What makes a measurement instrument valid and reliable?, *Injury*, **42**(3), pp. 236–240, 2011.
10. K. Topping, Peer assessment between students in colleges and universities, *Review of Educational Research,* **68**(3), pp. 249–276, 1998.
11. V. P. Kommula, J. Uziak and M. T. Oladiran, Peer and self-assessment in engineering students' group work, *World Transactions on Engineering and Technology Education*, **8**(1), pp. 56–60, 2010.
12. N. Falchikov, Involving students in feedback and assessment: A report from the Assessment Strategies in Scottish Higher Education (ASSHE) project, in *Peer assessment in practice*, S. Brown, Ed. Scotland: Staff Education Development Association (SEDA), pp. 9–22, 1998.
13. S. Ahmady, N. Khajeali, F. Sharifi and Z. Mirmoghtadaei, Educational intervention to improve preclinical academic performance: A systematic review, *Journal of Education and Health Promotion*, **8**, p. 83, 2019.
14. R. P. Machera, Teaching Intervention Strategies That Enhance Learning in Higher Education, *Universal Journal of Educational Research*, **5**(5), pp. 733–743, 2017.
15. C.-C. Foong and E. G. Daniel, Incompetent grounds in science students' arguments: What is amiss in the argumentation process?, *Procedia–Social and Behavioral Sciences,* **9**, pp. 1198–1207, 2010.
16. C. M. Brooks and J. L. Ammons, Free riding in group projects and the effects of timing, frequency, and specificity of criteria in peer assessments, *Journal of Education for Business*, **78**(5), pp. 268–272, 2003.
17. K. Willey and A. Gardner, Investigating the capacity of self and peer assessment activities to engage students and promote learning, *European Journal of Engineering Education*, **35**(4), pp. 429–443, 2010.
18. P. M. Sadler and E. Good, The impact of self-and peer-grading on student learning, *Educational Assessment*, **11**(1), pp. 1–31, 2006.
19. J. A. Ross, The reliability, validity, and utility of self-assessment, *Practical Assessment, Research, and Evaluation,* **11**(1), p. 10, 2006.
20. M. A. Jackson, A. Tran, M. P. Wenderoth and J. H. Doherty, Peer vs. self-grading of practice exams: Which is better?, *CBE – Life Sciences Education*, **17**(3), p. es44, 2018.
21. S. Quinton and T. Smallbone, Feeding forward: Using feedback to promote student reflection and learning – a teaching model, *Innovations in Education and Teaching International*, **47**(1), pp. 125–135, 2010.
22. R. M. Ryan and E. L. Deci, Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being, *American Psychologist*, **55**(1), p. 68, 2000.
23. E. L. Deci and R. M. Ryan, Self-determination theory: A macrotheory of human notivation, development, and health, (in English), *Canadian Psychology*, **49**(3), pp. 182–185, 2008.
24. C. Johns, *Becoming a Reflective Practitioner*, 3rd ed ed. Oxford: Wiley-Blackwell, 2009.
25. D. Boud, R. Keogh and D. Walker, *Reflection: Turning experience into learning*, London: Kogan Page, 1985.

26. P. Black and D. Wiliam, Inside the black box: Raising standards through classroom assessment, *Phi Delta Kappan*, **92**(1), pp. 81–90, 2010.
27. J. C. G. de Sande and J. I. Godino-Llorente, Peer assessment and self-assessment: Effective learning tools in higher education, *International Journal of Engineering Education*, **30**(3), pp. 711–721, 2014.
28. Y. Cinar and A. Bilgin, Peer assessment for undergraduate teamwork projects in petroleum engineering, *International Journal of Engineering Education*, **27**(2), pp. 310–322, 2011.
29. P. R. Pintrich, The role of metacognitive knowledge in learning, teaching, and assessing, *Theory into Practice,* **41**(4), pp. 219–225, 2002.
30. S. E. Carr and P. H. Johnson, Does self reflection and insight correlate with academic performance in medical students?, *BMC Medical Education*, **13**(1), p. 113, 2013.
31. R. M. Felder, G. N. Felder and E. J. Dietz, The effects of personality type on engineering student performance and attitudes, *Journal of Engineering Education,* **91**(1), pp. 3–17, 2002.
32. G. K. Tippin, K. D. Lafreniere and S. Page, Student perception of academic grading: Personality, academic orientation, and effort, *Active Learning in Higher Education*, **13**(1), pp. 51–61, 2012.
33. H. Verkade and R. J. Bryson-Richardson, Student acceptance and application of peer assessment in a final year genetics undergraduate oral presentation, *Journal of Peer Learning*, **6**(1), p. 2, 2013.
34. L. F. Siow, Students' perceptions on self-and peer assessment in enhancing learning experience, *Malaysian Online Journal of Educational Sciences,* **3**(2), pp. 21–35, 2018.
35. C.-C. Foong and E. G. Daniel, Students' argumentation skills across two socio-scientific issues in a Confucian classroom: Is transfer possible?, *International Journal of Science Education*, **35**(14), pp. 2331–2355, 2013.
36. M. Minkov, A revision of Hofstede's individualism-collectivism dimension: Old evidence and new data from 56 countries, *Cross Cultural & Strategic Management*, **24**(3), pp. 386–404, 2017.
37. C. Y. R. Loh and T. C. Teo, Understanding Asian students learning styles, cultural influence and learning strategies, *Journal of Education & Social Policy*, **7**(1), pp. 194–210, 2017.
38. E. S. C. Ho, Self-related cognition and mathematics performance: What we learned from PISA 2003 to PISA 2012, in *What we learned from PISA: The outstanding performance of students in Hong Kong and East Asia*, E. S. C. Ho, Ed. Singapore World Scientific, pp. 235–262, 2017.
39. D. L. Paulhus, J. H. Duncan and M. S. Yik, Patterns of shyness in East-Asian and European-heritage students, *Journal of Research in Personality*, **36**(5), pp. 442–462, 2002.
40. H. C. Boucher, Understanding Western-East Asian differences and similarities in self-enhancement, *Social and Personality Psychology Compass,* **4**(5), pp. 304–317, 2010.
41. C. C. Foong and Z. N. Anuar, Peer assessment, in *Enhancing student learning in medical schools*, J. H. Sim, C. C. Foong, W. H. Hong, V. Pallath and J. Vadivelu, Eds. Kuala Lumpur, Malaysia: University of Malaya Press, pp. 73–82, 2018.
42. N. M. Webb and A. S. Palincsar, Group processes in the classroom, in *Handbook of educational psychology*, D. C. Berliner and R. C. Calfee, Eds. USA: Prentice Hall International, pp. 841–873, 1996.
43. D. W. Johnson and R. T. Johnson, *Cooperation and competition: Theory and research*, Minnesota: Interaction Book Company, 1989.
44. A. Abdullah and P. B. Pederson, *Understanding multicultural Malaysia: Delights, puzzles and irritations*, Selangor, Malaysia: Prentice Hall, 2003.
45. S. M. Sim, C. C. Foong, C. H. Tan, P. Lai, S. S. Chua and M. Mohazmi, The use of jigsaw learning technique in teaching medical students prescribing skills, *Med. Teach.*, **36**(2), p. 182, 2014.
46. M. A. Alsubaie, Hidden curriculum as one of current issue of curriculum, *Journal of Education and Practice*, **6**(33), pp. 125–128, 2015.
47. C. Roberts, C. Jorm, S. Gentilcore and J. Crossley, Peer assessment of professional behaviours in problem-based learning groups, *Medical Education*, **51**(4), pp. 390–400, 2017.
48. R. Molina-Carmona, R. Satorre Cuerda, P. Compañ and F. Llorens Largo, Metrics for estimating validity, reliability and bias in peer assessment, *International Journal of Engineering Education*, **34**(3), pp. 968–980, 2018.
49. D. Statman, Modesty, pride and realistic self-assessment, *The Philosophical Quarterly*, **42**(169), pp. 420–438, 1992.
50. L. R. Harris and G. T. L. Brown, Opportunities and obstacles to consider when using peer-and self-assessment to improve student learning: Case studies into teachers' implementation, *Teaching and Teacher Education*, **36**, pp. 101–111, 2013.
51. A. M. Helmer, N. A. Slater, K. F. Marlowe, D. W. Surry and E. K. McCoy, Comparing faculty evaluations of student journal club presentations with student self-and peer evaluations during advanced pharmacy practice experiences, *Currents in Pharmacy Teaching and Learning*, **12**, pp. 564–569, 2020.
52. S. K. Lanning, T. H. Brickhouse, J. C. Gunsolley, S. L. Ranson and R. M. Willett, Communication skills instruction: an analysis of self, peer-group, student instructors and faculty assessment, *Patient Education and Counseling*, **83**(2), pp. 145–151, 2011.
53. J. R. Kogan, How to evaluate and give feedback, in *The academic medicine handbook*, L. W. Roberts, Ed. New York: Springer, pp. 91–101, 2013.
54. N. A. K. A. Holder, Z. L. Sim, C. C. Foong and V. Pallath, Developing a Reflection Guiding Tool for underperforming medical students: An action research project, *Tuning Journal for Higher Education*, **7**(1), pp. 115–163, 2019.
55. J. W. Creswell, *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*, Boston Pearson, 2012.

Chan Choong Foong is a senior lecturer at the University of Malaya, Malaysia. He received his PhD degree from the University of Malaya in the field of science education. He is a medical and physics education specialist, and he is currently exploring engineering education. His research interests include learning theories, student underperformance and support, and programme evaluation.

Peng Yen Liew is a senior lecturer at the Universiti Teknologi Malaysia (UTM) Kuala Lumpur, Malaysia. He received his PhD degree from the Universiti Teknologi Malaysia in the field of chemical engineering on process system engineering and process integration. He has four years of experience in coordinating and lecturing on the Chemical Process Plant Design Project for final-year students enrolled in Bachelor of Chemical Process Engineering. His research interest in engineering education concentrates on project-based learning in chemical engineering programmes.