JUTLP
Journal of University Teaching & Learning Practice

Developmental Changes of the Responses to Peer Assessment of Group Projects: A Case Study of Civil Engineering Undergraduates

Piyal Wasantha Pallewala Liyanage[a], Kirsten Black[b], and Sam Fragomeni[a]

[a] Victoria University, Australia; [b] RMIT University, Australia

Abstract

The changes in the responding behaviour to peer assessments of group projects as students progress through their studies were explored to assist educators in designing peer assessment tools to suit students' maturity levels and better understand the students' responses to peer assessments. We collected and analysed the responses to a peer assessment tool after group project assignments of a group of civil engineering undergraduate students in their second and third years. The responses displayed a relatively higher satisfaction level by the students on their peers' contribution to the group projects in both units, with a markedly higher satisfaction level in the third year of study than in the second year. The data also suggests that smaller groups tend to have higher peer satisfaction levels among group members than in bigger groups. Average peer rating versus group mark comparison was identified as a reliable indicator to identify groups that had internal issues while working on group projects but were not openly disclosed. Responses to individual criteria of the peer assessment tool confirmed that the group members are highly satisfied with each other's attendance in group meetings and are increasingly satisfied with each other's intellectual contribution towards the group project as the year level progresses. Given the participants in this study belong to one specific discipline, the results may not be generalisable to the entire student population, and we encourage more research in this area.

Editors

Section: Student Experience
Editor-in-Chief: Dr Joseph Crawford
Senior Editor: Dr Sally Ashton-Hay

Publication

Received: 01 April, 2023
Revision: 12 November, 2023
Accepted: 16 December, 2023
Published: 31 January, 2024

Citation

Pallewela Liyanage, P. W., Black, K., & Fragomeni, S. (2024). Developmental Changes of the Responses to Peer Assessment of Group Projects: A Case Study of Civil Engineering Undergraduates. *Journal of University Teaching and Learning Practice, 21*(1).

Introduction

Group projects have been used as a form of assessment in higher education for decades. Davies (2009) summarised a number of benefits of using group work as an assessment, including (1) the promotion of 'deep' and 'active' learning as opposed to 'surface' and 'passive' learning, respectively (Britton et al., 2017; Hammar Chiriac, 2014; Kremer & McGuinness, 1998; Lau et al., 2014; Michaelsen, 1992); (2) an expedient way of developing students' transferable skills (e.g. teamwork, communication, project management, adaptability and leadership) which are highly demanded by recruiters (Davies, 2009; Ravenscroft, 1997); and (3) providing a platform that facilitates the construction of knowledge and problem-based learning (Dolmans et al., 2001; Hendry et al., 1999).

Students, however, have expressed contrasting perceptions of group work assessments according to the literature. Daba et al. (2017) and Walker (2001) found, for example, that the majority of the students preferred group work over other types of assessments, whereas Knight (2004) reports that the students may instinctively prefer individual assessments, although they perform well in group assessments. Academics, on the other hand, see group work as an opportunity to ease the burden of marking and providing feedback (Freeman, 1995; Goldfinch & Raeside, 1990; Mantzioris & Kehrwald, 2013; Rafiq & Fullerton, 1996; Volkov & Volkov, 2015). Therefore, for academics/teachers, the benefit of group work is a pragmatic one.

Despite the advantages of group work assessments, some major issues have also been reported. Awarding the same mark to all group members with no appraisal of the potential differences of individual contributions has been a primary critique among others; poor and less-contributing students receive the same mark as those who are more competent and contributing in the same group. Actively obtaining the reward for no effort is generally referred to as 'free-riding'. Free riders cause the other motivational issue, the 'sucker effect', which arises as a result of the reduced input of more competent and willing students to a group project in response to the 'free riding' of other members (i.e. free riding themselves) (Davies, 2009; Elliott & Higgins, 2005; Heslop et al., 2017; Kerr, 1983; Lin et al., 2021; Ruël et al., 2003; Strong & Anderson, 1990). Goldfinch and Raeside (1990) proposed that having the students submit their individual portions of the project and/or a staff member sitting with the group to rate the contribution of individuals may counter those issues. However, Goldfinch and Raeside (1990) further commented on those suggestions, highlighting the drawbacks of either disrupting the spirit and not getting the full benefit of group work or being highly time-consuming for academics. Peer assessment, where the group members themselves assess each other's contribution, has been used as a relatively better approach to evaluating individual contributions in group projects (Conway et al., 1993; Goldfinch & Raeside, 1990; Ko, 2014; Sluijsmans et al., 2001; Stančić, 2021; Vaughan et al., 2019). The following sections discuss this approach in detail.

Peer assessment approach

Peer assessment, in some cases along with self-assessment, is an opportunity for students to quantitatively and/or qualitatively reflect on their strengths and weaknesses and to judge and compare each group member's contributions. This way, the students are also involved in the assessment process in a reasonably fair manner, as the group members should technically have the best sense of individual contributions. A litany of previous studies has reported positive impacts of peer assessment on students' motivation, participation and learning through group work. For example, Brindley and Scoffield (1998) found that the undergraduate students of their study felt that the peer assessment process had increased student motivation, the opportunity to compare and discuss the assignment, and the opportunity to gain knowledge and develop understanding. Also, Planas Lladó et al. (2014) observed students positively rating the peer assessment approach, highlighting that it motivates and facilitates their learning. Elliott and Higgins (2005) also report positive impacts of peer assessment on student motivation in group work based on the student responses.

Various peer assessment models have been used by various academics/tutors in higher education. While there is no universally accepted correct model fitting for all cohorts of students, the vast majority of the previously used models were inspired by the model proposed by Goldfinch and Raeside (1990) or its modified versions. Goldfinch and Raeside (1990) present their model in two parts. Part 1 is a questionnaire with a defined set of tasks related to the assessment, and each group member needs to write who contributed to each task. The number of times a particular student's name has been mentioned is used to calculate a 'Part 1 score' following equation 1. Part 2 questionnaire has a list of process skills as shown in Table 1 (considering a four-member group), and the students are required to award a mark between 0 and 4 to each group member. The 0-4 scale was defined by Goldfinch and Raeside (1990) as 0 – didn't contribute in this way, 1 – willing but not very successful, 2- average, 3- above average, and 4- outstanding. The individual marks are then summed up to obtain a score for each individual, and a 'Part 2 score' is then calculated following equation 2.

$$Part\ 1\ score = \frac{Number\ of\ mentions\ of\ individual}{Possible\ number\ of\ mentions} \qquad \text{Eq. (1)}$$

$$Part\ 2\ score = \frac{Actual\ sum\ scored}{Highest\ possible\ score} \qquad \text{Eq. (2)}$$

Table 1

Different criteria included in Part 2 of the peer assessment model proposed by Goldfinch and Raeside (1990)

Criterion	Member#1	Member#2	Member#3
1. Overall level of participation			
2. Understanding what was required			
3. Suggesting ideas			
4. Extracting something useful from the ideas			
5. Performing routine tasks			
6. Consolidating (i.e. drawing things together)			

7. Keeping the group going, particularly in difficult patches
8. Sorting out problems

A ‘PA factor’ is then calculated based on the Part 1 score and Part 2 score above, using equation 3.

$$PA\ score = \frac{1}{3}\ Part\ 1\ score + \frac{2}{3}\ Part\ 2\ score \qquad \text{Eq. (3)}$$

The weightings of 1/3 and 2/3 in equation 3 were based on the assumption that the contributions to the working of the group are roughly twice as important as ‘who did what’ (Goldfinch & Raeside, 1990).

Finally, the individual’s marks are calculated according to equation 4, after deriving a ‘PA factor’ by means of a ‘lookup table’ (Table 2). Interpolation was recommended by Goldfinch and Raeside (1990) for values in between those given PA scores in Table 2.

Table 2

PA factor lookup table suggested by Goldfinch and Raeside (1990).

PA score	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	1
PA factor (%)	0	20	40	60	80	100	110	120	130	140

$$Individual\ student's\ mark = (PA\ factor) \times (Group\ mark) \qquad \text{Eq. (4)}$$

Goldfinch and Raeside (1990) remarked that Part 1 of the model was helpful towards relieving the students’ anxiety on peer assessment, and students had confirmed it. However, many later developments of peer assessment tools either modified or did not heed it. For example, Rafiq and Fullerton (1996) identified project diaries, submitted along with the report, that assessors can use to allocate individual marks as a unique alternative to Part 1. Goldfinch (1994) also suggested that only Part 2 can be used based on the observation of the minor effect on the student responses after the omission of Part 1 scores, and the omission can result in halved administrative time for large classes. In addition, Conway et al. (1993) observed a poor correlation between Part 1 and Part 2 scores and a general reluctance of students to mention names in Part 1. They advocate a combined one-part assessment scheme and a modified calculation scheme to derive individual marks.

Peer assessment has been used with group work assessments in nearly all discipline domains, including science, engineering, art, mathematics and business studies. Computerised peer assessment tools have also been recently developed for improved efficiency of the assessment process and anonymously provide ratings and comments (e.g. Earnest et al., 2022; Havard et al., 2023; Kaewsaiha & Chanchalor, 2020; Kumar et al., 2019; Lin et al., 2021; Loureiro & Gomes, 2023).

Issues with the peer assessment approach

While peer assessment offers a sound means of allocating group marks to individual group members, a few critical issues with that approach have been discussed in the literature. Unfair, unreliable or biased marking has been reported as a serious concern by many researchers. Kennedy (2005) reports that peer ratings in most cases differ only marginally from an equal allocation, which was attributed to students being reluctant to be judgmental of their peers and the averaging process employed to determine the contributing factors. Mathews (1994) disagreed that the peer assessment approach is robust enough to be the sole source of informational input. Vu and Dall'Alba (2007) stress the requirement of further research to address the concerns related to the reliability of the peer assessment approach. Several reasons have been identified for the unfair marking of students, e.g. the general reluctance to judge the peers, particularly when a strong friendship/relationship has been established, and the presumption that downgrading the peers may result in friction among peers, including feelings of hurt and betrayal (Elliott & Higgins, 2005; Lin et al., 2021; Vu & Dall'Alba, 2007).

Reciprocity, which refers to the lack of fairness in rating due to the relationships between the rater and the other group members, was identified as the most worrisome issue and most difficult to counter by Magin (2001). A peer assessment experiment by Li (2001) found an inevitable bias in the students' rating, and a new normalisation process was introduced to iron out the inherent shortcomings of the approach. Collusion is another issue that can vitiate the fairness of the peer assessment approach. Song et al. (2017) identified two types of collusion – small-circle, which refers to the behaviours of students who form small circles and give higher peer ratings to each other, and pervasive, which refers to students assigning top ratings to all the submissions they review – and found that rating-based peer assessments are prone to have more collusion than ranking-based peer assessments. Rafiq and Fullerton (1996) also observed some potential cases of collusion in their case study. Collusion between some group members was identified as a possible reason for the observed very similar or even identical pattern of peer evaluations in Mathews (1994). One group member (or a few members) can be heavily penalised, or the marks can be heavily inflated due to collusion. Gender bias in ratings is another issue identified with the peer assessment approach. In groups with disproportionate numbers of males and females, same/opposite gender bias effects can be significant when the scores are averaged for individual group members, whereas those effects may cancel out in groups with a roughly equal number of males and females (Falchikov & Magin, 1997).

The maturity level of students can influence the characteristics of responses to peer assessments. For example, first-year undergraduate students may have a different perception and response behaviour to peer assessments than those of final-year undergraduate students. A generalised peer assessment approach to students of all year levels may not be fit for purpose and representative of reality. Accordingly, the success of the peer assessment approach in allocating group marks to individual group members can also be dependent on the maturity level of the raters, and this has not been thoroughly investigated in the literature to date. Understanding these behavioural changes in students enables educators to design and develop peer assessment tools to suit students' maturity levels for a more informed assessment of group projects. In addition, the inclusion of this additional dimension of peer assessment (i.e. the student maturity level) provides educators with an in-depth and rigorous understanding of the students' responses to peer assessment tools. The present study contributes to this by analysing the behavioural changes of the responses to peer assessment of second-year and third-year undergraduate students. More importantly, the responses were obtained from essentially the same group of students in their second year and third year to more realistically reflect on the developmental changes of their responses to peer assessments. The following sections detail the methodology and discuss the results.

Method

This case study was conducted at Victoria University (VU), Melbourne, Australia, using the students studying for the Bachelor of Civil Engineering course. VU is one of Australia's few dual-sector universities, currently enrolling over 40,000 students in higher education and vocational education and training (VU, 2021a). VU recently introduced the 'Block Model' of teaching and learning, where small, focused classes are conducted, and students study one unit at a time and complete it within four weeks (i.e. one block). The Block Model of teaching and learning at VU has shown effectiveness in improving the academic success of poorer-performing students (Klein et al., 2020).

Geomechanics (NEC2202) and Geotechnical Engineering (NEC3102) are two core units of the Bachelor of Civil Engineering course at VU that undergraduate students undertake during their second and third years of study, respectively. NEC2202 mainly focuses on the properties and engineering behaviours of geomaterials (i.e. soil and rock), whereas NEC3102 includes design applications related to soils and rocks, such as foundations and earth-retaining structures. In general, in both units, the relevant theories, concepts and practices are first introduced and discussed, followed by laboratory experiments, where applicable, and interactive real-world problem-solving exercises. The students who completed NEC2202 in 2020 and NEC3102 in 2021 were the participants of the study, who responded to the same peer assessment tool for three different group work assessments (one in NEC2202 and two in NEC3102). Both these units were delivered online in 2020 and 2021, and a total of 151 peer responses were used for the analysis.

Overview of the assessments and participants

Under the group work assessment of NEC2202, students were first asked to form groups of four by themselves. The lecturer then formed random groups for those who were left without a group. A total of 18 groups resulted in each having 2-4 members. More than 91% of the students were in groups of 3 or 4 members, meaning that only a few students formed 2-member groups. Each group had to work on a seepage flow analysis project worth 15% of the total marks using software simulations and theoretical calculations, after which each group submitted a detailed report (Note that due to the nature of the assessment, which had manual computations and software simulations of a groundwater seepage scenario that were interrelated, it was not possible to scale down the content of the assessment task to fit for the smaller sized groups). In addition to the group report submission, each individual student was asked to submit a completed peer assessment spreadsheet (detailed in the next section) to the online teaching platform 'VU Collaborate' within the appropriate submission folder. This separate online submission method for peer assessments was adopted to assure the students that their peers do not see their responses (i.e., anonymous). The authors postulate that maintaining the anonymity of peer assessment submission, as in this case, improves the response rate and the authenticity of responses. The students were pre-informed that the failure to submit the individual peer assessments would result in an automatic penalty of 5% of the final assessment mark, which helped contribute to a high response rate (~81%). In addition, the interest of students in the assessment might have motivated them to respond to the peer assessment. A total of 57 responses were collected from 70 students who completed this assessment in two block deliveries in 2020. All responses were complete and included in the analysis.

The same group of students progressed to the third year of the civil engineering course and completed the NEC3102 unit in 2021, except those who failed NEC2202 in 2020. These students completed two group work assessments under NEC3102, each worth 15% of the total marks of the unit. One assessment focused on a foundation design project, whereas the other was on the problem-based design of earth-retaining structures. Groups of both assessments comprised 2-4 members, with more than 87% of total students in 3-4 member groups. Due to the nature of the assessment, the foundation design project assessment was not scaled down to cater for smaller groups of 2, whereas the content of the problem-based design assessment was scaled down for these 2-member groups. The same peer assessment tool used in their second-year unit (NEC2202) was provided after each assessment, and the responses were collected. Each student was asked to submit the completed peer assessment spreadsheets individually to the 'VU Collaborate' online teaching platform, as per the case of NEC2202. A high response rate of ~83% was observed potentially, again attributed to the introduced 5% penalty of final assessment mark for no submission and the assurance of anonymity. A total of 94 responses were collected from 37 groups after both assessments in two block deliveries of the NEC3102 unit (a total of 104 responses were collected from 125 expected responses under both assessments, and 10 of them were incomplete, hence disqualified for analysis).

Peer assessment tool design

The design of the peer assessment tool used in this study was largely inspired by the peer assessment model proposed by (Goldfinch & Raeside, 1990). To maintain the simplicity of the peer assessment process, a one-part assessment approach was used, primarily based on significant recommendations in previous studies presented herein. This approach is comparable to part 2 of the Goldfinch and Raeside (1990) model. In effect, a simplified and user-friendly spreadsheet was created to collect the self and peer assessment marks out of 10 against six criteria, as shown in Table 3.

Table 3

The spreadsheet grading table used in the peer assessment

Assessment by	*'Your name'*

	Your grade – criterion 1	Your grade – criterion 2	Your grade – criterion 3	Your grade – criterion 4	Your grade – criterion 5	Your grade – criterion 6
'You'						
'Team member 2 name'						
'Team member 3 name'						
'Team member 4 name'						

The criteria and guidance for marking were provided separately, as shown in Figure 1. In this study, a broader rating scale of 1-10 for awarding marks was used compared to the 1-4 scale used in many previous studies. The proposed broader rating scale offers students greater granularity and rigour for assessing their peers more accurately and pragmatically. In addition, each criterion was elaborated with two extreme-end explanations (Figure 1). In being provided with these exemplified best (10/10) and worst scenarios (0/10) for each criterion, the students were given a clear understanding of the rating scale range to more confidently and accurately rate their peers. The criteria and extreme-end explanations used were previously applied in other units. This contrasts with the brief descriptions of different criteria (usually singly worded or a few words only) of peer assessments used in many previous studies, which can be ambiguous to the students.

A rather simple model that is used by Conway et al. (1993) was selected to allocate marks to individual group members based on peer assessments, as given in Equation 5.

$$Individual\ student's\ mark = (Individual\ weighting\ factor) \times (Group\ mark) \qquad \text{Eq. (5)}$$

The individual weighting factor is calculated by,

$$Individual\ weighting\ factor = \frac{Individual\ effort\ rating}{Average\ effort\ rating\ of\ the\ group} \qquad \text{Eq. (6)}$$

The discussion on the strengths and weaknesses of the selected model, using Equation 5, to allocate marks to individual group members is beyond the scope of this study.

Figure 1

Description of different criteria and the marking scale used in the peer assessment tool

Guide for the peer assessment (NOT to attach to the report):

10 marks	Marks awarded										0 marks
1. Regular attendance at group meetings											
Attended all meetings, stayed to agreed end, worked within timescale, active and attentive, prepared to be flexible about meeting times	10 ☐	9 ☐	8 ☐	7 ☐	6 ☐	5 ☐	4 ☐	3 ☐	2 ☐	1 ☐	Missed several/most meetings, always or often late, left early, digressed, gossiped most of the time
2. Contribution of ideas for the project											
Thought about the project in advance of meetings, provided workable ideas which were taken up by the group, built on others' suggestions, was prepared to test out ideas rather than keep quiet	10 ☐	9 ☐	8 ☐	7 ☐	6 ☐	5 ☐	4 ☐	3 ☐	2 ☐	1 ☐	Didn't come prepared, didn't contribute any ideas, tended to reject others' ideas without acceptable reasons
3. Researching, analysing and preparing material for the project											
Did what was agreed on, made an effort to understand background theory/material, did an equal share of the analysis, helped to evaluate the results	10 ☐	9 ☐	8 ☐	7 ☐	6 ☐	5 ☐	4 ☐	3 ☐	2 ☐	1 ☐	Didn't do the analysis, didn't do what was promised, didn't manage the workload, didn't get involved, allowed others to do the work
4. Contribution to cooperative group process											
Left personal differences outside the group, willing to review group progress and tackle conflict in the group, took on different roles as needed, kept group on track, willing and flexible but focused on task	10 ☐	9 ☐	8 ☐	7 ☐	6 ☐	5 ☐	4 ☐	3 ☐	2 ☐	1 ☐	Took no initiative, waited to be told what to do, always took the same role regardless of circumstance, created conflicts, were not prepared to review group progress
5. Supporting and encouraging group members											
Keen to listen to others, encouraged participation, enabled a collaborative learning environment, sensitive to issues affecting group members, supported group members with special needs	10 ☐	9 ☐	8 ☐	7 ☐	6 ☐	5 ☐	4 ☐	3 ☐	2 ☐	1 ☐	Sought only to complete the project, spoke over others, ignored their opinions, kept ideas to self only, insensitive to individual's needs, did not contribute to the learning process
6. Practical contribution to end product											
Willing to try new things, not hogging the tasks, made a high-level contribution, took own initiative, reliable, continued contribution till the end	10 ☐	9 ☐	8 ☐	7 ☐	6 ☐	5 ☐	4 ☐	3 ☐	2 ☐	1 ☐	Not willing to take on any task, did not take any responsibility, was unreliable so others felt the need to keep checking up, made a limited and poor quality contribution

Results and Discussion

The student responses to the different criteria of the peer assessment spreadsheet (as provided in Figure 1) for both NEC2202 and NEC3102 units were collated and analysed to derive meaningful and useful conclusions below.

Ideal responses

An ideal response is defined here as the response with a 10/10 marking for all six criteria of the peer assessment tool. The ideal response can be genuine, i.e., the rater is fully satisfied with each peer contribution, or it can result from biased ratings, as described in Section 1.2. In this study, 43.86% and 56.38% of the responses were identified as ideal responses in NEC2202 and NEC3102 units, respectively. Interestingly, both units show a noticeably higher percentage of ideal responses in comparison to many previous studies (e.g. (Goldfinch, 1994; Magin, 2001). While the issues such as bias, reciprocity, and collusion may have played a role in the displayed higher satisfaction of students about their student peer contribution towards the group project, the authors believe the mode of teaching and learning employed at VU (i.e., Block Model), which is not the case in the discussed previous studies, has also contributed to this. Typical class attendance is around 40% at most universities that utilise traditional teaching and learning environments, where data shows an attendance rate above 80% at VU under the Block Model of teaching and learning (Klein et al., 2020; VU, 2020). This markedly higher attendance rate is also indicative of students engaging with group work assessments more often, resulting in students rarely missing group meetings that, in turn, leads to better satisfaction among group members on each other's contribution (this effect is further discussed later in this section). In addition, the Block Model of teaching and learning has caused a significant rise in overall student academic performance. Loton et al. (2020) states that VU students are predicted to achieve over ten marks higher under the Block Model than in the traditional mode of teaching and learning. VU (2021b) also reports a jump of 7.9% overall pass rate of first-year students with the same assessments and a 13.9% increase in distinctions and higher distinctions at VU with the introduction of the Block Model in 2018, compared to the previous year. An increase in student learning, pass rate and average grades of first-year students after introducing the Block Model of teaching and learning is also reported by McCluskey et al. (2019). These remarkable improvements in student learning and academic performance would potentially have made a positive impact on group project assessments where the majority of the group members are able to deliver the required intellectual contributions towards the project to mutual satisfaction among group members.

It was further discovered that a marked increase in ideal responses occurred when students progressed from the second year to the third year of their study. Reasons behind this increase can include improved group cohesion and mutual understanding among group members as the students progress through the course. Lin (2018) also reports an improvement in friendship and learning networks among civil engineering students with time. However, it was noted that only 9% of the groups were unchanged across all three assessments under the two units. This change in group membership is a result of a) students not enrolling in the same block for NEC3102 as in the NEC2202, and b) more ‘like’ members combining to form groups after getting to know each other over an extended time period. In effect, the increased student maturity level and improved friendships/relationships within the groups are significant factors that lead to the higher number of ideal responses. As the groups had 2-4 members in the assessment tasks in both units, this study compared the percentages of ideal responses for the different group sizes, as shown in Figure 2.

Figure 2

Percentages of ideal responses for different group size

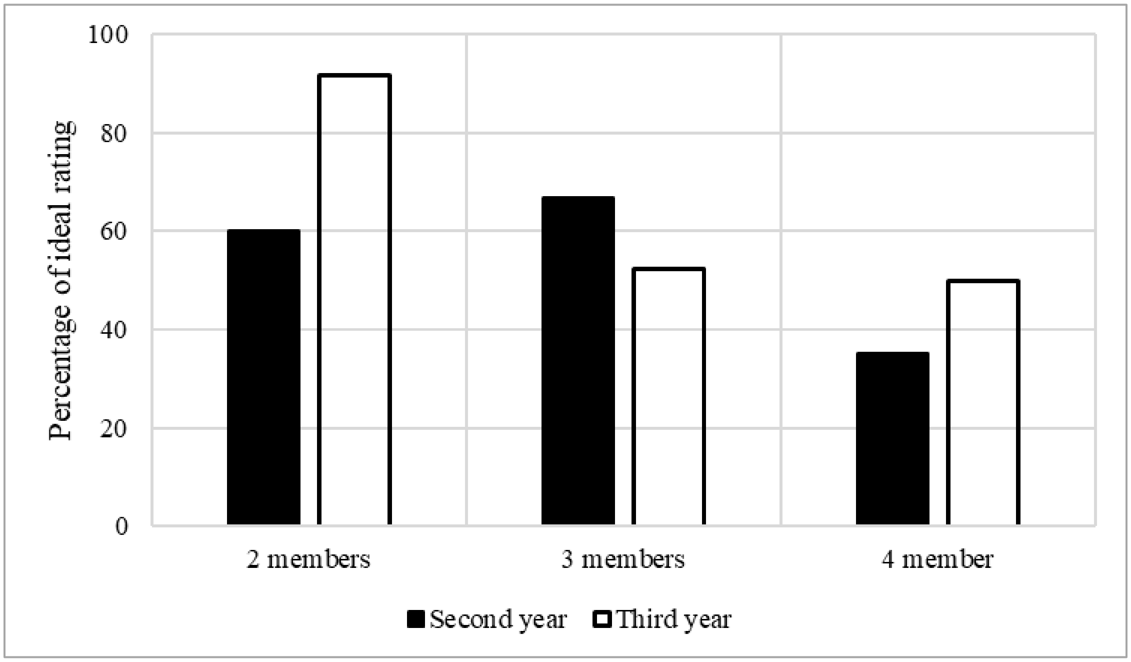


Figure 2 shows that the smaller groups (i.e., 2-3 members) are more likely to have a higher percentage of ideal responses than that of 4-member groups, irrespective of the students' maturity as the year level progresses. It is understandable that the higher the number of members in a group, the greater the chance of having more internal disagreements, sometimes leading to group disputes and issues that result in reduced peer ratings. With fewer members in a group, it is relatively easier to develop group cohesion and mutual understanding, leading to better mutual satisfaction. On the other hand, the anonymity of peer responses is weakened in smaller groups, particularly in 2-member groups where each group member is able to qualitatively deduce the other members rating with a great deal of certainty. This will lead group members to rate the other higher, irrespective of the contribution, to avoid possible post-conflicts. In addition, smaller groups are more susceptible to reciprocity issues that might have played a role here. Quantifying the relative contribution of each of these factors was not possible with the data available. Therefore, despite the relatively smaller number of 2-member groups, Figure 2 depicts some important, plausible characteristics of the group size effect on peer satisfaction levels of group assessments. However, the authors emphasise the importance of students working in relatively bigger groups than 2-3 member groups that provides them with an opportunity to improve their working in a team soft skill.

Peer rating versus group performance

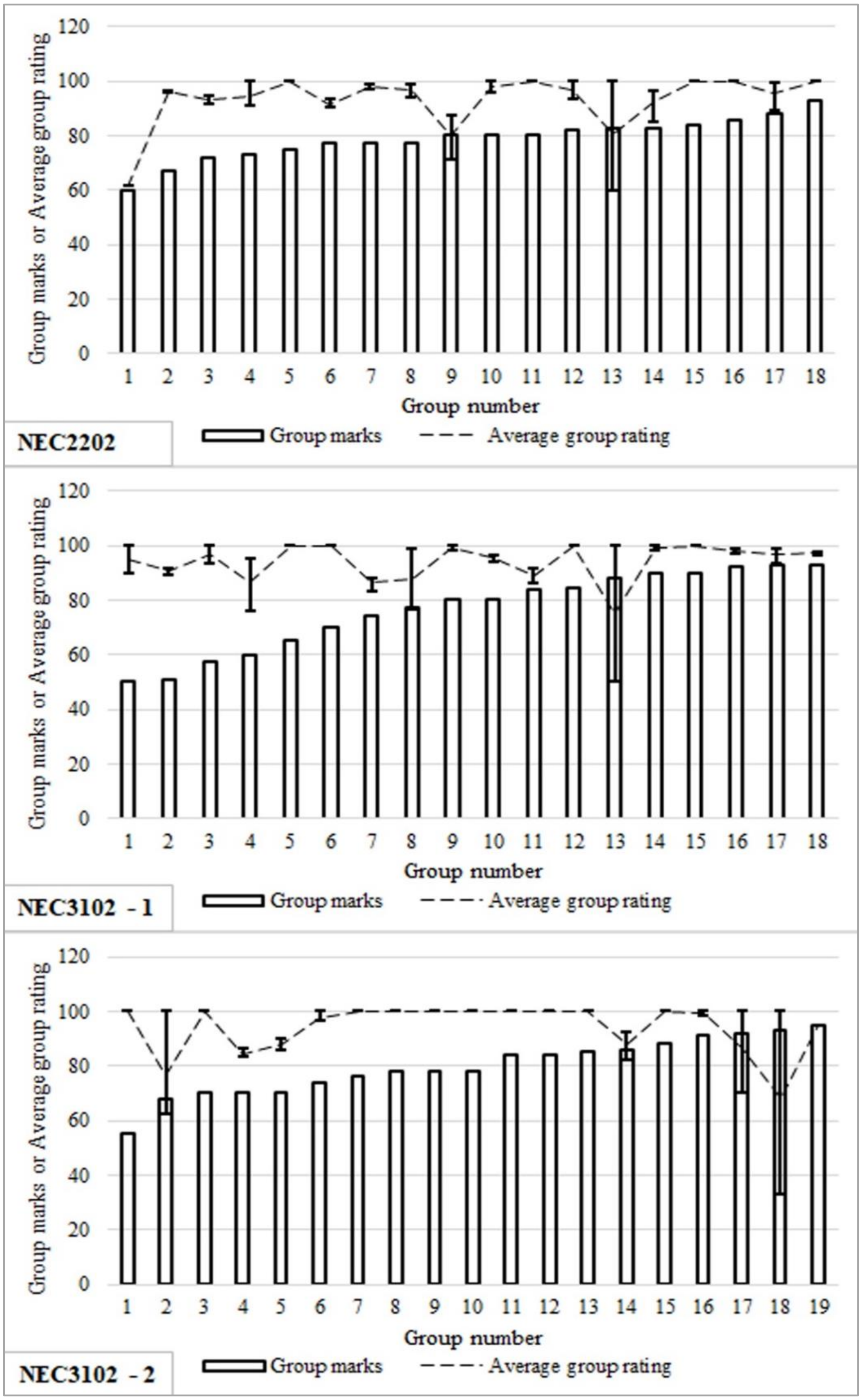
Peer assessment ratings can be viewed as a reflection of group synergy and conduct throughout the assessment period; a higher average group rating (i.e. the average of the marks from all members of the group) generally indicates the group's smooth conduct, which should lead to better performance in the assessment (i.e. the group mark) and vice versa. However, the possibility of a higher group mark with a lower average group peer rating is not ruled out. It can be an indication that one or a few members of the group have contributed to all or the majority of the assessment tasks to achieve higher marks and subsequently rated the other members with relatively low marks. It should also be noted that issues such as bias, reciprocity and collusion can complicate the understanding of the real group conduct during the project period using peer assessments.

The average group rating and group marks for all groups in the three assessments are plotted in Figure 3. No specific trends are apparent from Figure 3 for any assessment. It can be seen that the average group rating fluctuates throughout the range of group marks for NEC2202. A relatively higher number of groups have provided 100% average group rating (i.e. ideal response from all group members) in NEC3102 compared to that of NEC2202. A few cases of lower average group rating at lower group marks (<70%) can be observed in all three assessments of Figure 3. These can be examples of poor group conduct where the members did not support each other and performed as a group, which eventually led to poor peer ratings and group marks.

Interestingly, some groups that scored higher group marks (>80%) have also shown less average group rating but with a broader range of ratings, as indicated by the error bars in Figure 3. This may suggest that one or a few group members (possibly the more competent and keen members) had done the majority of the work to score higher marks and were not satisfied with their peers' contribution. In general, students rarely report the issues within groups, and no formal or informal complaints were reported to the academic in these cases. Therefore, the comparison of group performance with peer ratings is a reasonably robust way of understanding the group conduct and synergy, giving the teaching academic(s) a chance to follow up and discuss potential matters with individuals or the group, leading to improvements for future assessments. This is particularly important as students are generally reluctant to report or discuss issues with academics, especially in their initial years of study.

Figure 3

Group marks and average group ratings for the three group project assessments



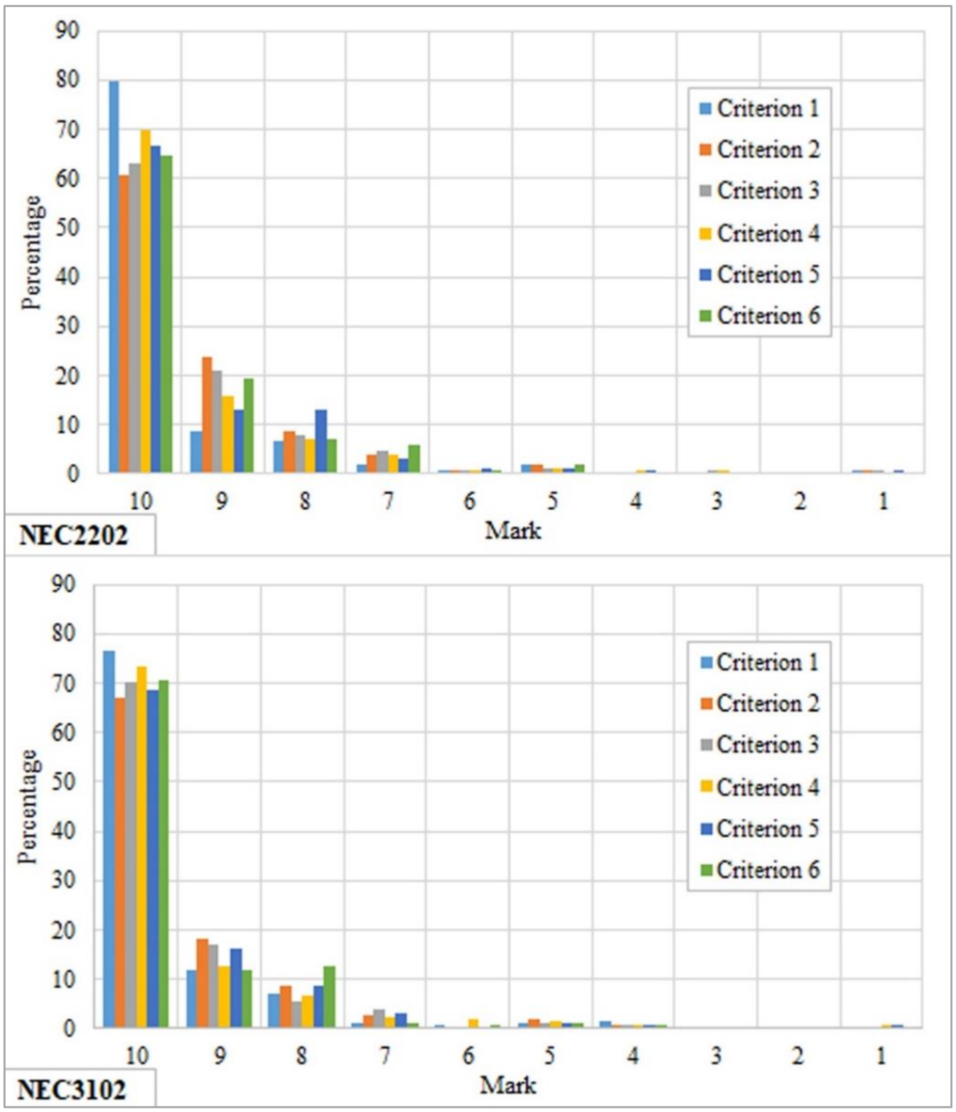
Responses to different assessment criteria

The student responding behaviour to the six different criteria of the peer assessment tool (as highlighted in Figure 1) is important to understand the common behaviours and pressing concerns that groups had that lead to respective ratings. This kind of analysis is also helpful for further improvements of the peer assessment tool.

Figure 4 shows the percentages of the marks given to each criterion in the two units considered (NEC2202 and NEC3102).

Figure 4

Percentages of responses received for different marks for each criterion of the peer assessment tool used

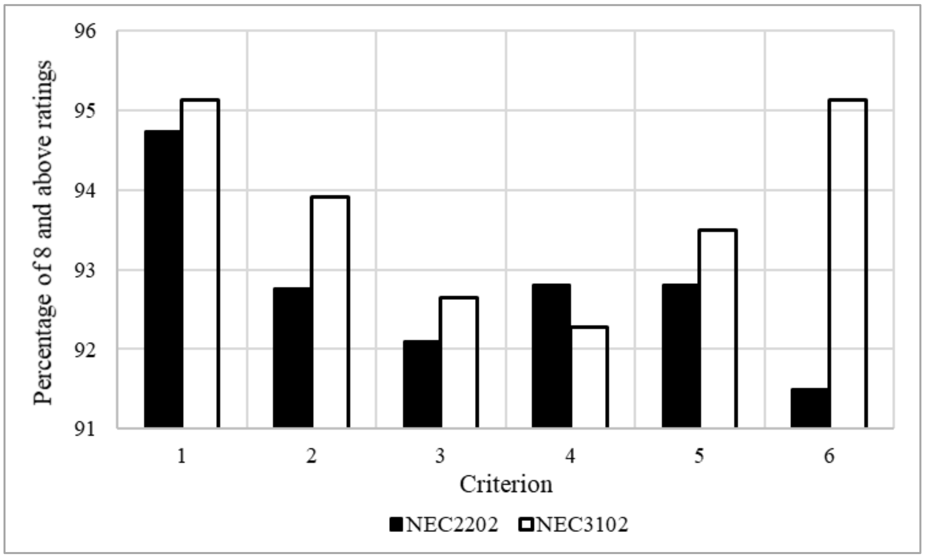


The greatest percentage of responses have been received for the maximum marks of 10 for all criteria, irrespective of the year level of the students. Criterion 1 (i.e. regular attendance at group meetings) received the highest percentage of 10/10 rating, followed by Criterion 4 (i.e. contribution to cooperative group process) in both units. This suggests that regular attendance at group meetings has been very good overall, and reasonably good cooperation between group members existed. The better attendance at group meetings can be correlated with the current teaching and learning approach at Victoria University, i.e., the Block Model as described earlier in this section; the attendance of students in classes has remarkably improved since the inception of the Block Model at Victoria University (Klein et al., 2020; VU, 2020), and generally students working on their group work assessments on the days they attend classes. In 2020 and 2021, these two units were remotely delivered, and most students organised their group meetings on virtual platforms, which offered them more flexibility to schedule their meetings. This may have also contributed to the higher satisfaction of students with their peers' attendance at group meetings. The second-highest percentage of 10/10 marks was on the contribution to cooperative group work, and this can also be correlated with the fact that most group meetings took place virtually. Students informally commented that they take their contribution more seriously, and the willingness to contribute, participate and be responsible is higher when the meetings are virtual.

The criterion with the least 10/10 mark is the contribution of ideas for the project/assessment (i.e. criterion 2). This is a common concern of group projects among students that not all members endeavour to contribute with ideas due to various reasons. Interestingly, the percentage increased from ~60% for NEC2202 to ~67% for NEC3102. This indicates that students are becoming more technically knowledgeable to provide more ideas and input for group discussions as they progress through the year level. However, it should be noted that criterion 2 has the highest percentage of 9/10 marks for both units, indicating that the contribution of ideas for the assessment by group members was generally good. When considering higher marks (8 or more), as shown in Figure 5, NEC3102 shows higher accumulated percentages for all criteria except for criterion 4 when compared to NEC2202. Although the difference is not significant for criteria 1-5, the result for criterion 6 (i.e. practical contribution to end product) is striking. Criterion 6 evaluates the practical contribution to the end product, and Figure 5 shows a greater satisfaction for the third-year unit (i.e. NEC3102) than that of the second-year unit (i.e. NEC2202). This can again be attributed to the improved technical competence of the students as they progress through year levels, enabling them to make more meaningful technical contributions to the project outcomes.

Figure 5

Percentage of 8 or above marks for different criteria of the peer assessment tool used



Conclusions

Peer assessment is a useful technique to allocate group marks to individual group members. This study explored the changes in the patterns of student responses to peer assessment criteria as the students progress through year levels. The peer assessment data collected after three group projects of civil engineering students at Victoria University in their second and third year of the study were analysed. In general, the peer ratings of students showed high satisfaction levels on peer contribution, which was attributed partly to the 'Block Model' of teaching and learning practice adopted at Victoria University, where students' attendance and academic performances have evidently improved after its introduction in 2018, with some potential issues with peer assessments such as bias, collusion and reciprocity. Another observation suggests that smaller groups tend to have a higher level of peer satisfaction than larger groups.

The trends of average group rating versus group marks were observed to be a strong indicator of the state of group conduct and synergy during group functioning. Group members showed a high level of satisfaction with their peers' attendance at group meetings irrespective of the year level, and the students are contributing more intellectually to group projects as their maturity level grows with progression through the year level of study. These observations were also correlated with the higher attendance and improved academic performances of students, which was evident after introducing the 'Block Model' of teaching and learning at Victoria University. Both these units were remotely delivered, and students used virtual platforms to have their group meetings. This offered them greater flexibility to work together as a team, leading to higher satisfaction with each other's attendance and contribution.

Although a sufficiently large data set was used in the analysis, the results of this study may lack generalisability when extrapolating to the entire student population due to the data source being limited to one specific cohort of students studying civil engineering at a single university. In addition, the students studied the units under the block model of learning and teaching; hence, the results of this study may deviate from those of other learning settings. Therefore, further research with other student samples is encouraged to test the propositions of this study.

Conflict of Interest

The authors disclose that they have no actual or perceived conflicts of interest. The authors disclose that they have not received any funding for this manuscript beyond resourcing for academic time at their respective universities. The authors report no usage of artificial intelligence in the design or development of this manuscript.

References

Brindley, C., & Scoffield, S. (1998). Peer assessment in undergraduate programmes. *Teaching in Higher Education*, *3*(1), 79-90.

Britton, E., Simper, N., Leger, A., & Stephenson, J. (2017). Assessing teamwork in undergraduate education: a measurement tool to evaluate individual teamwork skills. *Assessment & Evaluation in Higher Education*, *42*(3), 378-397. https://doi.org/10.1080/02602938.2015.1116497

Conway, R., Kember, D., Sivan, A., & Wu, M. (1993). Peer assessment of an individual's contribution to a group project. *Assessment & Evaluation in Higher Education*, *18*(1), 45-56.

Daba, T. M., Ejersa, S. J., & Aliyi, S. (2017). Student perception on group work and group assignments in classroom teaching: the case of Bule Hora University second year biology students, South Ethiopia--an action research. *Educational Research and Reviews*, *12*(17), 860-866.

Davies, W. M. (2009). Groupwork as a form of assessment: common problems and recommended solutions. *Higher Education*, *58*(4), 563-584. https://doi.org/10.1007/s10734-009-9216-y

Dolmans, D. H., Wolfhagen, I. H., Van Der Vleuten, C. P., & Wijnen, W. H. (2001). Solving problems with group work in problem-based learning: hold on to the philosophy. *Medical Education*, *35*(9), 884-889.

Earnest, M., Madigosky, W. S., Yamashita, T., & Hanson, J. L. (2022). Validity evidence for using an online peer-assessment tool (CATME) to assess individual contributions to interprofessional student teamwork in a longitudinal team-based learning course. *Journal of Interprofessional Care*, *36*(6), 923-931. https://doi.org/10.1080/13561820.2022.2040962

Elliott, N., & Higgins, A. (2005). Self and peer assessment–does it make a difference to student group work? *Nurse Education in Practice*, *5*(1), 40-48.

Falchikov, N., & Magin, D. (1997). Detecting gender bias in peer marking of students' group process work. *Assessment & Evaluation in Higher Education*, *22*(4), 385-396. https://doi.org/10.1080/0260293970220403

Freeman, M. (1995). Peer assessment by groups of group work. *Assessment & Evaluation in Higher Education*, *20*(3), 289-300.

Goldfinch, J. (1994). Further developments in peer assessment of group projects. *Assessment & Evaluation in Higher Education*, *19*(1), 29-35.

Goldfinch, J., & Raeside, R. (1990). Development of a peer assessment technique for obtaining individual marks on a group project. *Assessment and Evaluation in Higher Education*, *15*(3), 210-231.

Hammar C. E. (2014). Group work as an incentive for learning – students' experiences of group work. *Frontiers in Psychology*, *5.* https://doi.org/10.3389/fpsyg.2014.00558

Havard, B., Podsiad, M., & Valaitis, K. (2023). Peer assessment collaboration evaluation: an innovative assessment tool for online learning environments. *TechTrends*, *67*(2), 331-341. https://doi.org/10.1007/s11528-022-00832-8

Hendry, G. D., Frommer, M., & Walker, R. A. (1999). Constructivism and problem-based learning. *Journal of further and higher education*, *23*(3), 369-371.

Heslop, B., Stojanovski, E., Iveson, S., Paul, J., & Bailey, K. (2017). Respondent disengagement from a peer assessment instrument measuring Collaboration Viability. *Australasian Journal of Engineering Education*, *22*(2), 95-106. https://doi.org/10.1080/22054952.2018.1437676

Kaewsaiha, P., & Chanchalor, S. (2020). The development of an online peer assessment tool using google applications. *Proceedings of the 4th International Conference on Digital Technology in Education, Busan, Republic of Korea*. https://doi.org/10.1145/3429630.3429638

Kennedy, G. J. (2005). Peer-assessment in group projects: is it worth it? *Proceedings of the 7th Australasian Conference on Computing Education*, *42*, 59-65.

Kerr, N. L. (1983). Motivation losses in small groups: a social dilemma analysis. *Journal of Personality and Social Psychology*, *45*(4), 819.

Klein, R., Kelly, K., Sinnayah, P., & Winchester, M. (2020). The VU way: The effect of intensive block mode teaching on repeating students. *International Journal of Innovation in Science and Mathematics Education*, *27*(9).

Knight, J. (2004). Comparison of student perception and performance in individual and group assessments in practical classes. *Journal of Geography in Higher Education*, *28*(1), 63-81.

Ko, S. S. (2014). Peer assessment in group projects accounting for assessor reliability by an iterative method. *Teaching in Higher Education*, *19*(3), 301-314. https://doi.org/10.1080/13562517.2013.860110

Kremer, J., & McGuinness, C. (1998). Cutting the cord: student-led discussion groups in higher education. *Education + Training*, *40*(2), 44-49. https://doi.org/10.1108/00400919810206848

Kumar, K., Sharma, B. N., Nusair, S., & Khan, G. J. (2019). Anonymous online peer assessment in an undergraduate course: An analysis of students' perceptions and attitudes in the South Pacific. *2019 IEEE International Conference on Engineering, Technology and Education (TALE)*, *Yogyakarta, Indonesia, 2019*, pp. 1-8, https://.org/10.1109/TALE48000.2019.9226008

Lau, P., Kwong, T., Chong, K., & Wong, E. (2014). Developing students' teamwork skills in a cooperative learning project. *International Journal for Lesson and Learning Studies*, *3*(1), 80-99. https://doi.org/10.1108/IJLLS-03-2013-0018

Li, L. K. (2001). Some refinements on peer assessment of group projects. *Assessment & Evaluation in Higher Education*, *26*(1), 5-18.

Lin, J. W., Tsai, C. W., Hsu, C. C., & Chang, L. C. (2021). Peer assessment with group awareness tools and effects on project-based learning. *Interactive Learning Environments*, *29*(4), 583-599.

Lin, S. C. (2018). Evolution of civil engineering students' friendship and learning networks. *Journal of Professional Issues in Engineering Education and Practice*, *144*(4), 04018013. https://doi.org/doi:10.1061/(ASCE)EI.1943-5541.0000390

Loton, D., Stein, C., Parker, P., & Weaven, M. (2020). Introducing block mode to first-year university students: a natural experiment on satisfaction and performance. *Studies in Higher Education*, 1-24.

Loureiro, P., & Gomes, M. J. (2023). Online peer assessment for learning: findings from higher education students. *Education Sciences*, *13*(3).

Magin, D. (2001). Reciprocity as a source of bias in multiple peer assessment of group work. *Studies in Higher Education*, *26*(1), 53-63.

Mantzioris, E., & Kehrwald, B. (2013). Allocation of tertiary students for group work: methods and consequences. *Ergo*, *3*(2), 3-14.

Mathews, B. P. (1994). Assessing individual contributions: experience of peer evaluation in major group projects. *British Journal of Educational Technology*, *25*(1), 19-28.

McCluskey, T., Weldon, J., & Smallridge, A. (2019). Rebuilding the first-year experience, one block at a time. *Student Success*, *10*(1), 1-15. https://search.informit.org/doi/10.3316/informit.592416706809663

Michaelsen, L. K. (1992). Team learning: a comprehensive approach for harnessing the power of small groups in higher education. *To Improve the Academy*, *11*(1), 107-122.

Planas L. A., Soley, L. F., Fraguell S. R. M., Pujolras, G. A., Planella, J. P., Roura-Pascual, N., Suñol M. J. J., & Moreno, L. M. (2014). Student perceptions of peer assessment: an interdisciplinary study. *Assessment & Evaluation in Higher Education*, *39*(5), 592-610.

Rafiq, Y., & Fullerton, H. (1996). Peer assessment of group projects in civil engineering. *Assessment and Evaluation in Higher Education*, *21*(1), 69-81.

Ravenscroft, S. P. (1997). In support of cooperative learning. *Issues in Accounting Education*, *12*(1), 187.

Ruël, G. C., Bastiaans, N., & Nauta, A. (2003). *Free-riding and team performance in project education*. University of Groningen.

Sluijsmans, D. M., Moerkerke, G., Van Merrienboer, J. J., & Dochy, F. J. (2001). Peer assessment in problem-based learning. *Studies in Educational Evaluation*, *27*(2), 153-173.

Song, Y., Hu, Z., & Gehringer, E. F. (2017). Collusion in educational peer assessment: how much do we need to worry about it? *IEEE Frontiers in Education Conference (FIE)*, pp. 1-8.

Stančić, M. (2021). Peer assessment as a learning and self-assessment tool: a look inside the black box. *Assessment & Evaluation in Higher Education*, *46*(6), 852-864. https://doi.org/10.1080/02602938.2020.1828267

Strong, J. T., & Anderson, R. E. (1990). Free-riding in group projects: control mechanisms and preliminary data. *Journal of marketing education*, *12*(2), 61-67.

Vaughan, B., Yoxall, J., & Grace, S. (2019). Peer assessment of teamwork in group projects: evaluation of a rubric. *Issues in Educational Research*, *29*(3), 961-978. https://search.informit.org/doi/10.3316/aeipt.225410

Volkov, A., & Volkov, M. (2015). Teamwork benefits in tertiary education. *Education + Training*, *57*(3), 262-278. https://doi.org/10.1108/ET-02-2013-0025

Victoria University. (2020). New data shows student success has lifted again under VU's revolutionary block model. Retrieved 26/07/2021 from https://www.vu.edu.au/about-vu/news-events/news/new-data-shows-student-success-has-lifted-again-under-vu-s-revolutionary-block-model

Victoria University. (2021a). University Profile. Retrieved 26/07/2021 from https://www.vu.edu.au/about-vu/university-profile

Victoria University. (2021b). What the vu block model means for the future of education. Retrieved 26/07/2021 from https://www.vu.edu.au/about-vu/news-events/study-space/what-the-vu-block-model-means-for-the-future-of-education

Vu, T. T., & Dall'Alba, G. (2007). Students' experience of peer assessment in a professional course. *Assessment & Evaluation in Higher Education*, *32*(5), 541-556.

Walker, A. (2001). British psychology students' perceptions of group-work and peer assessment. *Psychology Learning & Teaching*, *1*(1), 28-36.