Online-based Rubric for Peer Assessment: Effectiveness and Implications

Kok Yueh Lee¹, David Hassell², Shana Mat Salleh³ & Thulasimani Munohsamy⁴

^{1,3,4}Universiti Teknologi Brunei, Brunei Darussalam & ²University of Bath, United Kingdom

ARTICLE INFO

Article history:

Received May 30, 2021

Revised June 30, 2021

Accepted July 02, 2021

Keywords:

Peer assessment
Online assessment
Higher education
Rubrics
Public speaking

Conflict of Interest:

None

Funding:

None

ABSTRACT

This study investigates using technology to promote authentic and meaningful learning in applying a peer assessment rubric for a public speaking assessment in a higher education institution in Brunei Darussalam. Three hundred six undergraduates from Universiti Teknologi Brunei's Schools of Business, Computing, and the Engineering Faculty conducted the assessments in real-time using online-based rubrics accessible via their smartphones or laptops. Comparisons were made between the lecturers' marks and students for each rubric criterion, and a set of questionnaires was distributed to investigate students' perceptions toward the peer assessment after the assessment. The results indicated a variable discrepancy between assessments by the lecturers and students for the rubric criteria. While in some disciplines, peer marking was found to overmark compared to the lecturer by more than 15%, in other cases, the marks were similar. Comparison between peer and lecturer assessment indicated that the level of agreement was sensitive to the lecturer, but less so between student cohort when assessed by the same lecturer. When differences were observed, there was no apparent discrepancy in an agreement between aspects of the rubric which evaluated content or delivery. Students' feedback revealed a positive response towards peer assessment but highlighted issues surrounding the technological aspects of the implementation process.

Corresponding Author: Kok Yueh Lee, Centre for Communication, Teaching and Learning, Universiti Teknologi Brunei, Tungku Highway, Gadong, BE1410. E-mail: kokyueh.lee@utb.edu.bn

© Kok Yueh Lee, David Hassell, Shana Mat Salleh, Thulasimani Munohsamy

This is an open access article under the CC BY-SA 4.0 international license.

1. Introduction

Assessments in classrooms have shifted dramatically from the more conventional approach such as tests or examinations to newer forms of assessments such as portfolios or the use of technology and inclusion of students as part of the assessment exercise. Assessment and feedback play a vital role in teaching and learning, and peer assessment is used as an additional means to strengthen this significant role of assessment (De Grez *et al.*, 2010). Students' involvement in assessment has been recognized as an essential aspect of learning and teaching, usually in the form of self or peer assessment, where students take on the assessor's role. Peer assessment is defined as "an arrangement in which individuals consider the amount, level, value, worth, quality, or success of the products or outcomes of learning of peers of similar status" (Topping, 1998:250). In peer assessment, students evaluate the work of their peers using a predetermined set of criteria or guidelines, and students assess their own work in self-assessment. A peer assessment component in a course assessment could promote students' participation, be accountable towards the teaching and learning process, increase attainment of skills and knowledge, and obtain feedback (Weaver and Cottrell, 1986). Therefore, peer assessment plays a significant role in informative assessment by involving the students to evaluate their peers, and with proper implementation, peer assessment can be used as a part of summative assessment. Peer assessment is not only evaluating the learning outcomes, but it is also a learning process (White, 2009).

Incorporating technology in peer assessments can help enhance students' engagement and promote self-learning (Kim and Bonk, 2006). However, some educators have reported setbacks in their attempts to integrate technology into the learning process regarding finding the balance between actual learning and technological distractions (Mouza and Lavigne, 2012). Educators must be clear on the feedback expectations and their justifications to encourage a form of self-regulated learning where students can see their goals and work towards achieving those goals (Steffens, 2006). It can be achieved by introducing a rubric assessment system to the students before the assessment process.

Henriksen et al. (2016) define a rubric assessment system as "*a matrix of explicit criteria used to evaluate student performance*". The implementation of rubrics as part of assessment has been an increasingly popular practice in education across all levels and various disciplines such as the humanities, business, social sciences, natural and applied sciences (Reddy & Andrade, 2010). Assessment rubrics can then be used to provide qualitative feedback and ensure both lecturers and students focus on what is required within the learning outcomes (Henriksen et al., 2016).

This research aimed to focus on using technology in the application of an online-based peer assessment rubric and evaluate the effectiveness of the rubric in a public speaking assessment in a higher education institution in Brunei Darussalam. Expanding on the preliminary work of Hassell and Lee (2020), the online technique and rubric were used with multiple lecturers for student cohorts studying on a range of courses within the University. In addition, the students conducted peer assessments of their speaking skills using an online-based rubric. The use of a technological component in the assessment also ensured a degree of anonymity on the student assessor, a factor that may have played a positive sign towards their participation in the exercise.

1.1. Objectives of the study

This study seeks to investigate:

- 1) how technology can be used to promote authentic and meaningful learning in applying a peer assessment rubric alongside evaluating the effectiveness of the rubric for a public speaking assessment in a higher education institution in Brunei Darussalam.
- 2) Concordance between the marks of the lecturers and students for each rubric criterion and feedback from the students after the assessments.

1.2. Significance of the study

Peer assessment can provide opportunities for students to assess each other's progress and subsequent learning. Utilizing the internet and technology as an alternative to basic paper assessments represents a shift away from the conventional style of assessment and grading system to one that incorporates and assimilates current assessment trends to engage and motivate the students to be involved in assessing their peers. Davies (2000) suggests that the combination of instruction through the internet can improve students' absorption learning, and in relation, peer assessment places the onus on the students to be responsible for their learning and the learning of others. Furthermore, it indirectly teaches the students to engage in thoughtful criticism to gauge their peers (Falchikov & Goldfinch, 2000) and thus helps them for the future by learning to provide and receive constructive criticism, a skill much needed in the workplace.

By addressing the potential differences between the grades awarded by peers and the lecturers, it is hoped that the information garnered could be used to improve our grading system further and by association with our rubrics criterion and guidelines.

2. Literature Review

Students' involvement in their presentation assessment is beneficial for their own learning (Cheng and Warren, 2005). However, the reliability and validity of the assessment and the students' attitudes towards peer assessment have always been questioned. De Grez et al. (2010) investigated the reliability and validity of peer assessments of oral presentation skills adopting a previously designed rubric. The rubrics assessed nine different criteria, which focused on content, delivery, and overall criterion. A total of 95 engineering students with a collection of 1105 oral presentations were assessed in their study, investigating the relationship between personal characteristics and both performance and assessment. The results showed that psychometric qualities are acceptable and that self-efficacy is an essential component in the assessment exercise. The study also revealed that the rubric assessment demonstrated consistency and validity in line with the assessment criteria on content and showed significant differences between the groups of students assessed; however, the variance was low.

De Silva's (2014) study investigated the impact of rubrics and self-assessment on ESL students' writing and speech competencies focusing on a government secondary school in Sri Lanka. Prior to the assessment, students completed a set of questionnaires, and a small sample of students and teachers were interviewed to investigate their perceptions of the use and implementation of rubrics. The students were divided into two groups, and the performances of both groups were recorded and compared. In the first group, the students did not know the rubrics, whereas the second group knew the rubrics and the criteria. The findings revealed that the second group performed significantly better than the first group, likely due to their knowledge of the rubrics' expectations and the explanations of the rubric criteria. Furthermore, when the students' assessments were compared with those of the lecturers, it was found that the second group had a high correlation of marks with the lecturers as opposed to the first group, who showed significant differences between their assessments and the lecturers' assessments. It indicated that students must be properly introduced to the rubric and be trained and guided in peer assessment to bring a positive outcome in terms of the teaching, learning, and assessment process.

Ballantyne et al. (2002) investigated peer assessment in large classes at the University of Technology, Australia. Nine hundred thirty-nine students from various classes completed a set of questionnaires to assess the students' views towards the implementation of peer assessment. They concluded that while there were difficulties in using peer assessment with large classes, the learning benefits of involving students in peer assessment outweigh the drawbacks. The results also indicated that the students were encouraged by the peer assessment as it promoted self-evaluation and self-reflection and provided them with the opportunity to develop valuable skills for subsequent employment. However, students were not in favor of evaluating peer competency or peers' ability to evaluate, and feelings that peers were either lenient or strict markers. Students also expressed that they had doubts about the fairness of peer assessment, and a majority of the students felt that the exercise was too time-consuming.

Wen and Tsai (2006) used an online-based peer assessment to investigate university students' views in Taiwan. Their findings correspond to other similar studies that students generally favored peer assessment as it allowed them to compare their work to peers. Students, however, were less enthusiastic about the feedback provided by their peers, criticisms in particular, and expressed a lack of uncertainty in having to assess their peers. Nevertheless, Wen and Tsai (2006) concluded that university students were in support of peer assessment.

Past literature has shown that students demonstrated positive attitudes towards peer assessment. There is also evidence of positive responses to using rubrics by college tutors and university instructors. For example, Powell's (2001) study of assessment in film and television production courses found that the instructor had a positive attitude towards using rubrics as an objective basis for evaluation. Similarly, Campbell (2005), in developing a rubric-based e-marking tool, reported that the instructors who used the e-marking tool with multiple classes enabled them to mark or grade students' work more consistently, efficiently, and fairly. Reitmeier et al. (2004, p.8), using rubrics to evaluate the oral presentation, reported that rubrics somewhat transformed the assessment from "subjective observations to specific performances".

Studies have shown correlations of marking between peer assessment and tutor. For example, Hafner and Hafner (2003) examined the reliability between peer assessment and tutor grading in a biology course involving 107 students. The findings showed that the marking between the tutor and student peer assessment was consistent. Similar results were also obtained in Dunbar et al. (2006) in a foundational general education course evaluating speeches by 100 students. Using Ebel's intra-class correlation (consistency), the study showed high inter-coder reliability for the entire scale (.96) and each of the evaluation criteria (.82–.97). All in all, these studies show that the use of rubrics can lead to a relatively common interpretation of student performance (Reddy and Andrade, 2010).

Peer and self-assessments have been implemented in classroom assessments over the years, as illustrated in the literature. Generally, lecturers and students demonstrated positive attitudes towards the use of peer and self-assessments, but there was also a sense of mixed feelings towards both assessments. Lecturers and students acknowledge the usefulness of these assessments. However, there were still some reservations relating to the amount of time required and rater reliability. Undoubtedly, it is good to have various classroom teaching methods rather than just focusing on one form of assessment. With the introduction of self and peer assessment, lectures enable students to be responsible for their learning, promoting autonomous learning.

3. Method

3.1. Course

The study was conducted at a higher education institution in Brunei. English is the primary medium of instruction in all the schools in Brunei, and as such, the English proficiency and fluency level is considerably high.

Public Speaking, an element of communication skills, is one of the topics covered in the core module, Effective Communication, undertaken by all first-year students enrolled at the University. Public Speaking skills are delivered over two weeks out of fourteen, and a public speaking assessment was carried out subsequently. Speakers must speak for 3 – 5 minutes on a range of topics as pre-approved by tutors. Topics ranged from bioengineering, motivational talks, cosmetic surgery, and gender equality. The study was conducted in the first semester, from Weeks 11 - 14, 2017/18.

3.2. Lectures

The public speaking assessment was conducted by three lecturers, all non-native English speakers, delivering the module to students from different disciplines in the institution. The lecturers' background is shown in Table 1.

Table 1. Background of Lecturers

Lecturer	Number of teaching experience in Higher Education	Discipline taught
A	4 years	School of Computing and Informatics (SCI) Civil Engineering (CE)
B	11 years	School of Business (SB) Petroleum & Chemical Engineering (PCE)
C	16 years	Electrical and Electronic Engineering (EEE) School of Applied Sciences and Mathematics (SASM)

3.3. Students

Three hundred six undergraduates from various disciplinary backgrounds took part in the study. The students were all in their first year of study, pursuing their degree in Applied Sciences and Mathematics (N= 29), Business and Information Systems (N= 100), Computing and Informatics (N= 88), Civil Engineering (N= 21), Electrical and Electronic Engineering (N= 38), and Petroleum and Chemical Engineering (N= 30). The students conducted the peer assessments in real-time using online-based assessment rubrics, accessible through their laptops or mobile phones. The majority of the students were nationals of Brunei, all non-native speakers of English, with a few international students from Malaysia, Dubai, and The Gambia. After the peer assessment, all students were asked to complete an anonymized questionnaire based on a Likert seven-point scale (Likert, 1932) consisting of closed and open-ended questions to determine their attitudes towards self and peer assessment. Questions were developed based on analysis of past questionnaires used in the literature on rubrics, prior lecturer experience delivering peer and self-assessments, and the research questions outlined in the Introduction section. Table 2 shows the summary of research participants in the study; the total number of students, the number of students who took part in the peer assessment, and the overall response rate for the questionnaire.

Table 2. Background of Students and Response Rate

	Number of students	Peer assessment	Questionnaire response rate
Applied Sciences & Mathematics	29	29	10 (34%)
School of Business	100	100	54 (54%)
School of Computing & Informatics	88	88	71 (81%)

Civil Engineering	21	21	14 (67%)
Electrical & Electronic Engineering	38	38	12 (32%)
Petroleum & Chemical Engineering	30	30	6 (20%)

3.4. Rubric Assessment

The rubric used was designed to specifically highlight all the evaluation criteria that encompass an accomplished public speaker, assessing both the process of the assessment (delivery) as well as the product (speech content) (Montgomery, 2002). In addition, the descriptions within the rubric were carefully crafted to avoid any critical or emotional implications hidden or otherwise to avoid creating any undue bias from the assessors (Nilson, 2003).

Under delivery, the presenters would be assessed based on their appearance, body language, eye contact, language, voice, and pacing and timing. For content, the presenters are assessed based on their introduction, conclusion, purpose, and originality. The rubric was provided to the students at least one week before the assessment to allow for the students to study and realize the required criteria for them to reflect on their skills and, at the same time, prepare them to reflect on their peers and their skills. It was particularly significant as the students were first-time users of the rubric, so proper guidance was required to understand the rubric and the assessment requirements fully. The students would then assess their peers during their 3-to-5-minute presentations based on the criteria within the rubric.

3.5. Data Analysis

The peer assessment exercises were collected from the google form responses and exported into Microsoft Excel to be analyzed. The average and standard deviation were calculated and compared to determine the variation between different cohorts. Further analysis compared the average peer assessment mark for each student against that awarded by the lecturer for each section within the rubrics. As the individual peer assessment average marks are non-discrete numbers while the lecturer response is, comparisons are presented based on an agreement to within the same grade (i.e., lecturer awards 3, peers award $2.5 < x < 3.49$) and within plus or minus one grade (i.e., lecturer awards 3, peers award $1.5 < x < 4$). Correlation analysis of the results was performed using Spearman's correlation coefficient to identify qualitative trends within the data, and the results were taken to be statistically significant based on the data provided in Zar (1984) using $\alpha \leq 0.025$. A questionnaire was administered to the students after the peer assessment exercise. The questionnaire was completed anonymously and consisted of statements and an open-ended question to investigate students' views and perceptions on the rubrics-based peer assessment. The questionnaire was devised based on a seven-point Likert-type scale (Likert, 1932) with 1 = Strongly Disagree, 7 = Strongly Agree, and 4 = Neither Disagree or Agree. The strong agreement is viewed as the average means of responses from 5 to 7, i.e., the higher the average, the stronger the agreement. The similarity between responses was evaluated using Analysis of Variance (ANOVA) tests with a significance level (α) of 0.05

4. Results and Discussion

4.1. Correlation between lecturers and students' assessments

All 306 students across the disciplines mentioned above completed the peer assessment using the online google form. Figures 1 and 2 summarise the correlation of marks between lecturers and peer assessment across all disciplines.

Figure 1. Correlation of Marks between Lecturers and Peer Assessment in all the Disciplines (Percentage Similarity of Marks between Lecturers and Peers)

Figure 2. Correlation of Marks between Lecturers and Peer Assessment with Percentages of Marks within ± 1

Table 3 shows the mean score and standard deviation for each cohort for peer assessment and lecturer assessment. Results indicate that except for SASM, students over the mark when compared to the lecturer. The results also indicate that the mean marks awarded by each lecturer seem to be similar independent of the cohort being marked. However, there is a considerable difference in the mean between lecturer A and lecturers B and C. Table 4 provides the Spearman Rank correlation between the peer and lecturer mark for each cohort, and except for CE, all cohort marking exhibits a positive correlation between peer and lecturer ranking of students' ability within the cohort. It indicates that while peer assessment does not accurately capture quantitatively the students' ability perceived by each lecturer, it captures to a certain extent the qualitative order of student abilities within a cohort.

Table 3. Summary of Cohort Average and Standard Deviation (SD) for Peer and Lecturers Mark

	<i>n</i>	Peer assessment		Lecturer assessment	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Applied Sciences & Mathematics (C)	29	72.61	4.97	70.03	5.83
School of Business (B)	100	82.87	5.33	70.74	12.26
School of Computing & Informatics (A)	88	72.38	7.21	56.38	11.56
Civil Engineering (A)	21	76.57	5.07	56.90	10.67
Electrical & Electronic Engineering (C)	38	78.68	3.92	72.00	2.97
Petroleum & Chemical Engineering (C)	30	83.11	4.43	73.06	5.87

Table 4. Summary of Spearman Rank Coefficient Values for Marking between Peer and Lecturers Mark

	<i>n</i>	Spearman Rank Coefficient	Correlation?
Applied Sciences & Mathematics (C)	29	0.47**	YES
School of Business (B)	100	0.75***	YES
School of Computing & Informatics (A)	88	0.69***	YES
Civil Engineering (A)	21	-0.01	NO
Electrical & Electronic Engineering (C)	38	0.33*	YES
Petroleum & Chemical Engineering (C)	30	0.57***	YES

*. Correlation is significant at the 0.05 level (two-tailed)

**. Correlation is significant at the 0.02 level (two-tailed)

***. Correlation is significant at the 0.002 level (two-tailed)

Figures 1 and 2 show similarities and discrepancies in the marking between the lecturers and peers across the disciplines. BUS, SCI, and CE displayed the most significant incongruity in all the nine rubric criteria (as shown in Figure 1) as compared to other disciplines (EEE, PCE, and SASM), which in the latter case were all marked by lecturer C. All the students irrespective of cohort seemed to have a different interpretation and judgment on *appearance* when compared to the lecturers. There was relatively weak agreement amongst EEE, PCE, and SASM for *body language*, *eye contact*, and *language and voice*. The results seemed to indicate that students in EEE, SASM, and PCE were more inclined to grade similarly to that of the lecturer in content-related criteria (*Introduction*, *Conclusion*, *Central Ideas and Purpose*, and *Content and Originality*) the level of agreement was less with PCE. The results indicate that while there is some degree of bias in agreement based on student cohort, a more significant bias is observed based on the lecturer assessing the module, also seen from the mean results in Table 3. Staff interpretation of what is and is not acceptable, for instance, the cultural background, has a more significant bearing on agreement than the students themselves. Lecturer A is shown to have the lowest level of agreement with their students, and as both lecturers A and B are Bruneian, it is surmised that the educational background of the staff at higher levels has a more significant impact than their formative years in education.

The findings also show a high tendency for students to give their peers higher or lower grades than the lecturers in all the disciplines apart from CE in terms of *eye contact*, *language and voice*, and *pacing and timing*. There was also a disparity in the BUS and PCE cohorts' evaluation of *appearance* (50% and 48% agreement, respectively). These results contradict studies by Hafner and Hafner (2003) and Dunbar *et al.* (2006), where when the reliability between peer-grading and instructor grades were examined, they

concluded that the instructor and student ratings were remarkably consistent. Similar findings were also displayed in Davey and Palmer's (2012) study where the students' and lecturer marking were consistent, although; the students tend to overmark and undermark compared to the lecturer, an outcome similar in the present study. Dunbar *et al.* (2006) study showed high inter-coder reliability for the entire scale (.96) and each of the evaluation criteria (.82–.97), unlike the results found in this study.

4.2. Questionnaire Findings

Subsequent to the peer assessment, students completed an anonymous questionnaire based on the peer assessment exercise. As mentioned, a seven-point Likert scale was utilized, ranging from Strongly Disagree (1) to Strongly Agree (7), with Neither Agree or Disagree as neutral (4). The results from the questionnaire centered on three main themes: experiences, opinions, and understanding. Results of a one-way analysis of variance (ANOVA) test indicating the agreement between the questionnaire responses of each cohort to questions is presented in Appendix A, Table A1. While there is a great deal of variation in the p-values, in all cases, they are greater than 0.05, indicating that the null hypothesis is accepted and that all student cohort responses for each question are statistically similar.

4.2.1. Experiences of the students

Figure 3 shows the students' experiences in peer assessment. Students stated that they were critical in assessing their peers and thought that rubrics is a fair means of assessment. Students also felt that their peers were competent to participate in the peer assessment exercise and would like more assessments of the exact nature to be carried out in the future. It is worth noting that SASM students were not as exuberant towards conducting the peer assessment as students in the other departments, which could be behind the low response rates of the questionnaire (34.48%). On the other hand, PCE students were to some extent more in favor of evaluating their peers with rigor and were critical in assessing others despite the minimal response rate (20%).

Interestingly, while CE students found peer assessment to be a helpful activity and would like similar assessments to be carried out in the future, they were not particularly critical in assessing others but felt that their peers were competent to assess their speech (72.7% response rate). Generally, all the students across all disciplines found the peer assessments exercise beneficial and approached the exercise seriously. The results of this study are similar to previous studies by Ballantyne *et al.* (2002) and Wen and Tsai (2006), where the university students had positive attitudes towards peer assessment. However, the group deprived of the

solution was indifferent towards peer assessment in the future.

Figure 3. Students' Experiences of the Peer Assessment

Figure 4 illustrates the students' opinions towards peer assessment. The opinions of students towards the Overall exercise vary. On the whole, all the students who were not keen to give feedback to their peers found the exercise a difficult process, particularly for EEE students. SASM students, in part favor, were not particularly in favour of peer assessment. There was generally a low agreement to lecturer feedback as well

as preference towards peer assessment, below 4 average respectively. However, there was broad agreement amongst students preferring lecturer assessment as compared to peer assessment. It corresponds with Nilson's (2003) findings in which most students were against peer assessment due to the heavy responsibility and the possibility of peer retaliation, preferring lecturer-based assessment. There also seemed to be some contradictions in students' opinions towards the assessment. On the one hand, they found the exercise challenging (broad agreement) on the other, the agreement level for preference not to conduct peer assessment on others was considerably low (average of 4.33), which could imply that students eventually

developed a rather positive attitude towards peer assessment a trend which was also demonstrated in students' experiences earlier, 'peer assessment is a worthwhile activity. Primarily, students in the present study were ambivalent towards peer assessment, similar to the study by Ballantyne et al. (2002), where students felt positive of the peer assessment; however, the students expressed concerns towards their own as well as peers' competency in assessing; issues of fairness, feelings that their peers were either being lenient or harsh markers and the majority of students felt peer assessment was too time-consuming.

Figure 4. Students' Opinion of the Peer Assessment

4.2.2. Students' understanding

Figure 5 shows the students' understanding of the peer assessment exercise. All students had a positive perspective of peer assessment. On the whole, there was broad agreement in all the disciplines where students felt that peer assessment was helpful in their learning (above 5.3 average) and it motivated them to learn (above 4.8 average). CE students (response rate 72.3%), notably, found the entire peer assessment exercise invaluable. SASM students, albeit positive, were less positive as compared to other disciplines, particularly in view of giving feedback to their peers, which they did not think significantly of (4.2 average). Through the peer assessment exercise, students acquired a better understanding of what was expected for the public speaking assessment (above 5.2 average) and what was expected of them in the assessment (above 5.4 average). Students also felt that they learned through the peer assessment exercise (above 4.8 average) and acquired more knowledge regarding the module using rubrics (above 5.0 average). Andrade and Du (2005) found similar findings that the rubrics help students to focus, produce higher quality work, acquire better grades, and feel less anxious about an assignment. The students also emphasized that grading is fair and transparent in assessment marking using rubrics.

Figure 5. Students' Understanding of the Peer Assessment

5. Conclusion

This paper presents a study investigating the implementation of technology during the delivery of a peer assessment exercise as part of an assessment in public speaking. The public speaking assessment was successfully conducted using an online rubric accessible via student smartphones or laptops and promoted instantaneous feedback and learning through applying the rubric to peers' work. Student feedback via questionnaire identified limitations to the technique, explicitly surrounding internet connection speed and connectivity, but overall, students were positive towards the experience and its impact on their learning. Despite some students identifying issues related to their poor understanding of the expectations and specifications of the rubric criterion, the use of technology provided an enhanced experience through the delivery of instantaneous feedback and peer-to-peer learning via the evaluation of positive and negative rubric criteria during public speaking. Comparison between lecturer and peer-assessment marks found variable discrepancies between lecturer and peer assessment marks depending on the student cohort and individual lecturer. These differences ranged up to fifteen percentages, with agreement found to be more sensitive to lecturer than student cohort. There was no apparent bias in observed discrepancies for specific aspects of the rubric based on whether they covered evaluated content or delivery, and these discrepancies make the technique unsuitable for summative assessment. Overall, the exercise was found to be a beneficial experience for student learning. However, given the quantitative discrepancies between marks awarded, future use of the technique should be limited to formative rather than summative assessment, where the provision of instantaneous feedback would provide an obvious benefit for student learning.

References

- Andrade, H. G. (2000). Using rubrics to promote thinking and learning. *Educational Leadership*, 57 (5), 13–19.
- Andrade, H. G., & Du, Y. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research & Evaluation*, 10 (3), 1–11.
- Ballantyne, R., Hughes, K., & Mylonas, A. (2002). Developing procedures for implementing peer assessment in large classes using an action research process. *Assessment & Evaluation in Higher Education*, 27 (5), 427–441.
- Campbell, A. (2005). Application of ICT and rubrics to the assessment process where professional judgment is involved: the features of an e-marking tool. *Assessment & Evaluation in Higher Education*, 30 (5), 529–537.
- Cheng, W., & Warren, M. (2005). Peer assessment of language proficiency. *Language Testing*, 22 (1), 93–121.
- Davey, K. R., & Palmer, E. (2012). Student Peer Assessment: A research study in a level III core course of the bachelor chemical engineering program. *Education for Chemical Engineers*, 7 (3), e85–e104.

- Davies, P. (2000). Computerized Peer Assessment. *Innovations in Education and Training International*, 37 (4), 346–355.
- Dunbar, N. E., Brooks, C. F., & Kubicka-Miller, T. (2006). Oral communication skills in higher education: Using a performance-based evaluation rubric to assess communication skills. *Innovative Higher Education*, 29 (4), 22–30.
- Falchikov, N., & Goldfinch, J. (2000). Student Peer Assessment in Higher Education: A Meta-Analysis Comparing Peer and Teacher Marks. *Review of Educational Research*, 70 (3), 287–322.
- De Grez, L., Valcke, M., & Berings. (2010). Student response system and learning oral presentation skills. *Procedia - Social and Behavioral Sciences*, 2 (2), 1786–1789.
- Hafner, J., & Hafner, P. (2003). Quantitative analysis of the rubric as an assessment tool: an empirical study of student peer-group rating. *International Journal of Science Education*, 25 (12), 1509–1528.
- Hassell, D. G., & Lee, K. Y. (2020). Evaluation of Multi Peer and Self-Assessment in Higher Education: A Brunei Case Study. *International Journal of Innovative Teaching and Learning in Higher Education*, 1 (1), 37–53.
- Henriksen, C. B., Bregnhøj, H., Rosthøj, S., Ceballos, A., Kaas, H., Harker-Schuch, I., Andersen, I., Larsen, S., & May, M. (2016). Technology enhanced peer learning and peer assessment. *Tidsskriftet Læring Og Medier (LOM)*, 9 (16). <https://doi.org/10.7146/lom.v9i16.24415>
- Kim, K. -J., & Bonk, C. J. (2006). The future of online teaching and learning in higher education. *Educause Quarterly*, 29 (4), 22–30.
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*.
- Montgomery, K. (2002). Authentic Tasks and Rubrics: Going Beyond Traditional Assessments in College Teaching. *College Teaching*, 50 (1), 34–40.
- Mouza, C., & Lavigne, N. (2012), *Emerging Technologies for the Classroom: A Learning Sciences Perspective*. Springer Science & Business Media, New York.
- Nilson, L. B. (2003). Improving student peer feedback. *College Teaching*, 5 (1), 34–38.
- Powell, T. A. (2001). *Improving Assessment and Evaluation Methods in Film and Television Production Courses*. Capella University, Minneapolis.
- Reddy, Y. M., & Andrade, H. G. (2010). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education*, 35 (4), 435–448.
- Reitmeier, C. A., Svendsen, L. K., & Vrchota, D. A. (2004). Improving oral communication skills of students in food science courses. *Journal of Food Science Education*, 3 (2), 15–20.
- De Silva, R. (2014). Rubrics for assessment: Their effects on ESL students' authentic task performance. *Center for English Language Communication 4th Symposium Proceedings*.
- Steffens, K. (2006). Self-regulated learning in technology-enhanced learning environments: Lessons of a European peer review. *European Journal of Education*, 41 (3–4), 353–379.
- Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research*, 68 (3), 249–276.
- Weaver, R. L. & Cottrell, H. W. (1986). Peer evaluation: A case study. *Innovative Higher Education*, 11 (1), 25–39.
- Wen, M. L., & Tsai, C. -C. (2006). University students' perceptions of and attitudes toward (online) peer assessment. *Higher Education*, 51 (1), 27–44.
- White, E. (2009). Student perspectives of peer assessment for learning in a public speaking course. *Asian EFL Journal*, 33 (1), 1–36.
- Zar, J. H. (1984), *Biostatistical Analysis* (2nd ed.). New Jersey: Prentice-Hall, Inc., Eaglewood Cliffs.