



Exploring usage of summative peer assessments in engineering education

J. Petrović¹

University of Zagreb Faculty of Electrical Engineering and Computing
Zagreb, Croatia
0000-0002-2335-0287

P. Pale

University of Zagreb Faculty of Electrical Engineering and Computing
Zagreb, Croatia
0000-0003-2171-7302

Conference Key Areas: *Assessment*

Keywords: *peer assessment; communication skills; peer evaluation; engineering education; peer grading; peer scoring*

ABSTRACT

Summative peer assessment is an assessment method where the one's work is typically graded by several other anonymous peers using predefined criteria. The value of summative peer assessments in higher education stems from the fact that they can provide scalability in assessment for large enrollment classes for a variety of different assessment types. The main disadvantages of using summative peer assessments are questionable validity and reliability. In this paper, the first results of using summative peer assessments in a large enrollment professional skills course at the University of Zagreb, Faculty of Electrical Engineering and Computing are reported and discussed. The main research question of this work is how well, given specific conditions of the conducted summative peer assessments, do assignment credits assigned by peers correlate with assignment credits assigned by course lecturers. Data were obtained from four summative peer assessments through the course. A random sample of 50 submitted works per peer assessment was evaluated by course lecturers and corresponding assignment credits were compared to assignment credits awarded by students. Data analysis results suggest a moderate to high correlation between several measures of assignment credits awarded by peers and lecturers.

¹ *Corresponding Author*

J. Petrović

Juraj.Petrovic@FER.hr



1 INTRODUCTION

Peer assessment is “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” [1]. Peer assessments can be either summative or formative in their nature [2]. The main goal of formative peer assessments is to provide the author of the submitted work with feedback to help them plan their learning [2], while the main goal of summative peer assessments, also known as *peer grading* [3] or *peer marking* [4], is to assign grades or assignment credits to the author of the assignment.

Research results suggest formative peer assessments, in which students typically get feedback on their assignments through textual comments has a positive influence on learning outcomes related to the assignment as well as on other factors including learning strategies, academic mindset [5], [6], and motivation for learning [7]. When using peer assessments in a summative context, the psychometric properties of validity and reliability of peer assessments are a major concern since they depend on students’ ability to produce valid and reliable evaluations of peers’ work. While students tend to perceive peer assessments as beneficial, some students express a dislike of awarding a grade to their peers [8]. A meta-analysis of peer assessment research revealed the average Pearson’s correlation coefficient between peer and teacher ratings of 0.63 [9]. Research has also indicated that quality of peer assessments’ psychometric properties can be improved by training and experience of peer assessors. Student attitudes towards peer assessment are also positively influenced by training and experience [10], although some authors suggest peer grading can undermine effects of peer feedback [11].

This short paper contributes to the body of related research by reporting the first results of a research exploring the correlation between teachers’ and students’ ratings on a summative peer assessment conducted among the first-year undergraduate engineering students in a professional skills course called *Communication skills*. The assignment credits assigned by students (peers) were compared with those assigned by course lecturers to investigate how valid peers’ credits are if they are used in summative context under conditions described in the paper.

2 METHODOLOGY

2.1 Context

This research was conducted in context of a first-year course *Communication skills* at the University of Zagreb, Faculty of Electrical Engineering and Computing. Around 700 students are enrolled in the course every year. Trough the semester, students have a total of six summative peer assessments in which the following assignments are assessed: (1) writing a formal email message according to an individualized scenario, (2) writing a cover letter and a narrative resume for a job application, (3) capturing a photograph, (4) capturing a video, (5) creating a slideshow presentation with narration, (6) creating a two-minute video presentation with audio narration and



subtitles. The assignment credits a student finally receives are based on peers' ratings of the submitted assignment (~65%). A student can achieve the remaining assignment credits (~35%) by submitting ratings of five peers' assignments. Following the advice in published research [12], students are, aside from a small amount of assignment credits (external motivation), provided with a clear and understandable evaluational framework and they have a guarantee of anonymity in the peer assessment process. To submit their ratings students use an online form (Moodle *Quiz* type activity) with a set of evaluation questions which are available to them in advance. Questions typically have four offered answers to each question, for example:

Is there a paragraph (part of text) that states who the e-mail sender is and what is their role? Is that part of the text clear?

- a. No, that part of the text does not exist.*
- b. Yes, that part of the text exists, but it is completely unclear.*
- c. Yes, that part of the text exists and is written somewhat clearly.*
- d. Yes, that part of the text exists and is written quite clearly.*

The evaluation form for each assignment contains questions that can roughly be divided into two categories: subjective questions like the above example, where the difference or exact criteria for answers b) and c) might vary between students, and objective questions like *Does the submitted work have subtitles?*, where the answers or answer criteria are close to indisputable. Answers to some of the objective questions (for example the presence of subtitles in a video file) are assessed automatically using software support. In order to discourage students from submitting invalid or unreliable assessments, students are warned that each of such attempts will result in losing all assignment credit for that complete assignment. This rule is enforced only on objective questions that are automatically removed from further processing.

2.2 Data

Data presented in this paper was collected during the academic year of 2020/2021 in four out of six summative peer assessments (all except (3) capturing a photograph, (4) capturing a video). In each of the peer assessments, ratings submitted by students are screened for inconsistencies in objective questions and assessments by their authors are removed from further processing. Additionally, about 5% to 10% of submitted works with highest standard deviation of assignment credits assigned through peer assessment are screened by lecturers. The assignment credits for each submitted work are afterwards awarded either by course lecturers or by averaging the assignment credits of the remaining peer assessments.

In order to investigate the relationship between assignment credits assigned by peers and those assigned by course lecturers, 50 randomly selected works for four peer assessments have been evaluated and assigned credits by both peers and course lecturers.

3 RESULTS

Results of the analysis of correlation between the lecturers' assignment credits and assignment credits awarded by by peers measured as (1) average value of peers' assignment credits, (2) average value of peers' assignment credits without minimal and maximal credits value, and (3) median value of peers' assignment credits, are presented in Table 1. The correlation coefficients (Pearson's r) are based on assessment of 50 submitted works for each peer assessment and are all significant above the $p = 0.001$ level. The average values (M) of lecturer's and peers' assignment credits are also given in the table for each of the peer assessments.

Table 1. Correlation between lecturers' assignment credits and other measures

Peer assessment	Lecturer's assignment credits	Peers' assignment credits	Peers' assignment credits without max and min	Median peers' assignment credits
(1) E-mail message	$M = 81.92\%$	$M = 85.63\%$ $r = 0.564$	$M = 86.52\%$ $r = 0.576$	$M = 86.60\%$ $r = 0.574$
(2) Resume and cover letter	$M = 91.52\%$	$M = 88.47\%$ $r = 0.813$	$M = 89.87\%$ $r = 0.811$	$M = 89.90\%$ $r = 0.812$
(3) Slideshow	$M = 89.12\%$	$M = 90.98\%$ $r = 0.686$	$M = 91.37\%$ $r = 0.708$	$M = 91.44\%$ $r = 0.705$
(5) Video presentation	$M = 88.45\%$	$M = 89.01\%$ $r = 0.812$	$M = 90.24\%$ $r = 0.775$	$M = 87.99\%$ $r = 0.824$

M – mean, r – Pearson's r

The overall results of peer assessments regarding the achieved assignment credits are relatively high, between 85.3% and 91.44% of the total assignment value. This can likely be attributed to the availability of evaluation criteria in advance helping students know what exactly is expected from them.

The three measures of calculating total assignment credits from peers' assignment credits have shown a strong correlation with assignment credits awarded by course lecturers. Overall, it seems that the average value of peers' assignment credits without maximal and minimal values best correlates with lecturers' assignment credits. The correlation values range from 0.564 to 0.824 indicating a strong correlation. The obtained results are encouraging regarding the potential of summative peer assessments in the described context. It is interesting that the correlation coefficients are the highest for the video presentation assignment, in which there are several objective questions, but also for the resume and cover letter assignment, where most evaluation questions are rather subjective in their nature.



4 CONCLUSIONS AND FUTURE WORK

In this paper the first results of a comparison between lecturers' and peers' assessments of students' work in a professional skills for engineers course are described. The obtained results suggest a moderate to strong correlation between peers' and lecturers' grades under given conditions and provide some support for claims about validity and reliability of this kind of assessments. Questions that remain and should be addressed in a follow-up in-depth investigation include how the relative ratio of subjective to objective questions affects the correlation of assignment credits, how high is the inter-rater reliability of course lecturers' assignment credits, and whether different assignment credit calculation models can be used, for example, to identify students whose evaluations are most similar to lecturers' and achieve a higher correlation by assigning a greater weight to their evaluations.

REFERENCES

- [1] Topping, K. (1998), Peer Assessment Between Students in Colleges and Universities, *Review of Educational Research*, Vol. 68, No. 3, pp. 249–276.
- [2] Topping, K. (2009), Peer Assessment, *Theory Into Practice*, Vol. 48, No. 1, pp. 20–27.
- [3] Freeman, S. and Parks, J. W. (2010), How Accurate Is Peer Grading?, *CBE Life Sciences Education*, Vol. 9, No. 4, pp. 482–488.
- [4] Khatoon, R. and Jones, E. (2021), Flipped small group classes and peer marking: incentives, student participation and performance in a quasi-experimental approach, *Assessment and Evaluation in Higher Education*, Vol. 0, No. 0, pp. 1–18.
- [5] Li, H., Bialo, J. A., Xiong, Y., Hunter, C. V. and Guo, X. (2021), The Effect of Peer Assessment on Non-Cognitive Outcomes: A Meta-Analysis, *Applied Measurement in Education*, Vol. 34, No. 3, pp. 179–203.
- [6] Li, H., Xiong, Y., Hunter, C. V., Guo, X. and Tywoniw, R. (2020), Does peer assessment promote student learning? A meta-analysis, *Assessment and Evaluation in Higher Education*, Vol. 45, No. 2, pp. 193–211.
- [7] Topping, K. (2005), Trends in Peer Learning, *Educational Psychology*, Vol. 25, No. 6, pp. 631–645.
- [8] Falchikov, N. (1995), Peer Feedback Marking: Developing Peer Assessment, *Innovations in Education and Training International*, Vol. 32, No. 2, pp. 175–187.
- [9] Li H. *et al.* (2016), Peer assessment in the digital age: a meta-analysis comparing peer and teacher ratings, *Assessment and Evaluation in Higher Education*, Vol. 41, No. 2, pp. 245–264.
- [10] van Zundert, M., Sluijsmans, D. and van Merriënboer, J. (2010), Effective peer assessment processes: Research findings and future directions, *Learning and Instruction*, Vol. 20, No. 4, pp. 270–279.
- [11] Liu, N. and Carless, D. (2006), Peer feedback: The learning element of peer assessment, *Teaching in Higher Education*, Vol. 11, pp. 279–290.
- [12] Andersson, M. and Weurlander, M. (2019), Peer review of laboratory reports for engineering students, *European Journal of Engineering Education*, Vol. 44, No. 3, pp. 417–428.