# Effect of self and peer assessments on mathematics learning achievement

**Baiduri**
University of Muhammadiyah Malang, Indonesia
✉ baiduri@umm.ac.id[*]

**Abstract**
Assessment is an essential part of the learning process in higher education. As a result, the assessment should involve students in addition to lecturers. The purpose of this study is to examine the self and peer assessments on the answers to midterm and final-semester exams given by lecturers and their relationship with student achievement. The participants in this study were 114 fifth-semester mathematics education students who took differential equations courses. Data were gathered from student assessment scores on the middle and final test and student achievement in the lecturer's differential equations courses. The data were analyzed using descriptive statistics and an independent sample t-test to describe the self and peer assessments. Pearson correlation analysis was also used to describe the relationship between self and peer assessment and learning achievement. The study revealed no significant difference between self and peer assessments, both midterm and final semester assessments. The midterm assessment said that the average self-assessment was higher than peers. Meanwhile, the assessment average scores obtained in the semester's final exam were nearly identical between the two assessors. There is a positive and significant correlation with a high level of self and peer assessments, both in the midterm and final semester, and a positive and significant correlation with a moderate level of self and peer assessments with learning achievement. These findings can be used as a basis for lecturers to consider when involving students in assessing student achievement.

## INTRODUCTION

Assessment is an unavoidable component of educational systems because it can influence learning and teaching by judging the level of achievement of learning outcomes. When done authentically, it provides feedback and revision to improve understanding (Al-Abdullatif, 2020; Meihami & Esfandiari, 2020). Furthermore, assessment can motivate students in the learning process in a meaningful way (Meihami & Esfandiari, 2020). Teachers regularly assess their students' learning because assessment is essential in teaching and learning. Teacher evaluations, sometimes called traditional evaluation systems, have some drawbacks (Meihami & Esfandiari, 2020). Teachers also make decisions regarding the quality of students' responses do not measure all aspects of student learning (Al-Abdullatif, 2020; Meihami & Esfandiari, 2020).

Based on the shortcomings of the teacher's assessment, alternative assessment techniques that are effective and influential must evaluate students' educational development. Alternative reviews have been found to benefit the instructional process because they include procedures and practices easily integrated into student activities (Al-Abdullatif, 2020; Iraji et al., 2016). Students' participation in the assessment process promotes critical skills such as responsibility, assessment, and independence. Participation in evaluations enables students to develop self-regulation skills

and engages them in becoming active participants in their learning (Hassell & Lee, 2019; Kilic, 2016). Self-assessment and peer-assessment are the two types of alternative assessment approach used to evaluate student achievement improvement (Al-Abdullatif, 2020; Meihami & Esfandiari, 2020; Oren, 2018). To obtain a holistic view of students' contributions, self-assessment and peer-assessment are frequently used with teacher assessment to arrive at a student's overall course grade (Alias et al., 2015). These encourage students' active participation in the assessment process by raising their awareness of the assessment criteria and improving and developing the quality of learning. Students excel as peer evaluators, with assignments that serve as high validity and reliability (Iglesias Pérez et al., 2020).

The process by which students evaluate their work concerning their performance, ability, or achievements using specific criteria is known as self-assessment (Golightly, 2021; Oren, 2018; Stančić, 2020). Students have good self-assessment skills, but they tend to underestimate their performance compared to teachers and peer assessment (Alias et al., 2015; Golightly, 2021). Self-assessment was sometimes problematic because students believed it was impossible to be objective when assessing their work (Lindblom-Ylänne et al., 2006). Peer assessment is a classroom activity in which students score, grade, or provide feedback on the work of their peers (Al-Abdullatif, 2020; Hassell & Lee, 2019; Iraji et al., 2016; Oren, 2018) or by scoring, commenting, and scoring and commenting (Chen et al., 2020). Peer assessment is exceptionally effective at improving students' understanding of the current task's requirements as well as their ability to plan steps to improve solutions to future unknown problems (Brignell et al., 2019), positive influence on learners' cognitive, metacognitive, and self-regulation abilities (Abrache et al., 2021; Omar. 2019), a successful strategy for encouraging students to compare and reflect on their peers' work (Chen et al., 2020), as well as enhances students' learning achievement, performance, self-efficacy, and critical thinking skill (Chang et al., 2020; Omar. 2019).

Self and peer-assessment are critical components of assessment for learning practice. Assessing their own or others' work can assist students in developing their understanding of learning objectives and success criteria. According to research, when students participate actively in their learning and assessment, they make more progress. In recent years, self and peer assessment practices have increased in various areas of higher education (Meihami & Esfandiari, 2020; Nawas, 2020; Oren, 2018). Both assessments encourage students to participate actively in their learning (Hassell & Lee, 2019). According to students, peer and self-assessment contribute to their learning by providing effective feedback, creating a supportive learning environment, and encouraging collaboration among students (Ndoye, 2017).

All assessments involve judgment; however, scoring systems in mathematics tend to reward quick and objective answers rather than mathematical reasoning, which is arguably more important but more difficult to judge (Burkhardt & Swan, 2012). Furthermore, there are four current scoring systems, which are as follows: point-based scoring, in which a numerical value is assigned to the method, accuracy, or explanation at each solution step; criteria-based scoring, in which all responses are graded using predefined descriptors; rubric-based scoring retains the holistic component of criteria-based assessment, but levels are assigned to different aspects of performance; and comparative assessment, in which responses are ranked by assigning relative ratings rather than ratings based on criteria (Brignell et al., 2019; Burkhardt & Swan, 2012). Self and peer assessments are used in this study to refer to point-based assessments. At each step, scores are assigned based on the accuracy and correctness of the answers. This research differs

14

from prior research, which used pairwise comparative assessment in developing test questions on the concept of multivariable calculus (Jones & Alcock, 2014).

Previous research on the use of self and peer assessments in the learning process has included studies in mechanical engineering (Hassell & Lee, 2019), language (Birjandi & Tamjid, 2012; Meihami & Esfandiari, 2020; Oren, 2018; Vasileiadou & Karadimitriou, 2021), geography (Golightly, 2021), and art (Chen et al., 2020). Similar research is still scarce in mathematics lessons in higher education associated with learning achievement. In this study, learning achievement is defined as the final grade given to students by the lecturer, including attendance components (10%), assignments (30%), mid-semester tests (30%), and final semester exams (30%). The aims of this research are 1) investigating the significance of differences between self-assessment and peer-assessment when assessing mid and final exams, 2) investigating the degree of relationship between self and peer assessment when assessing the middle and final exams, and 3) investigating the effect of self and peer assessment on learning achievement.

## METHODS

### Research Types and Approaches

This study aimed to describe the self and peers assessment and the impact on student achievement in Differential Equations Course. Thus, this study employed the descriptive statistical method and inference research with a quantitative approach.

### Research Procedure

This study examines the significance of self-assessment and peer-assessment and their influence on academic achievement at university. Participants in the study were 114 fifth-semester students from three classes of the mathematics education department, University of Muhammadiyah Malang, who took the differential equation course for the academic year 2020/2021, which researchers taught. There are 37 (32.46%) students in Class 5A, 38 (33 .33%) students in Class 5B, and 39 (334.21%) students in Class 5C. All participants have taken the mathematics learning evaluation course. The lecture process is carried out for 16 weeks effective days, with two midterm and final exams. Midterm and final semester exam questions are given in eight-week and sixteenth weeks, respectively. Lectures and exams are carried out online synchronously.

### Data Collection

The data were obtained from the assessment scores on the lecturer's midterm and final exam questions. The questions have fulfilled the criteria for content validity because they were used to assess the mastery of the materials that have been taught. The questions were derived from textbooks commonly used in differential equations courses to ensure consistency. For each exam, there were three questions. The time is given to students to carry out self, and peer assessment was two days after the examination. For the peer assessment to be objective, no two students judged each other. Suppose student A assessed student B's written answer; student B was not allowed to judge student A's. The students were not given the answer key or steps to solve the given questions but the maximum score obtained if the answer was correct. The scoring was based on the accuracy and correctness of the answers according to the assessors' perceptions.

### Data Analysis

Descriptive and inference statistics were used to analyze the data in this study. A descriptive statistical analysis and an independent sample t-test were used to describe the assessment

performed by the students and their peers. Pearson's correlation analysis was conducted to determine the correlation between assessments conducted by self and peers and their effect on learning outcomes. The Pearson r value represents the magnitude of the effect, where $< 0.1$ indicates a very small effect, 0.1-0.3 indicates a small effect, 0.3-0.5 indicates a moderate effect, and $> 0.5$ indicates a large effect (Goss-Sampson., 2020). JASP (Jeffrey's Amazing Statistics Program) version 0.14 software was used in all analyses.

## RESULTS AND DISCUSSION
### Description of Self and Peer Assessment Scores
This section discusses the description of self and peer assessment as a result of descriptive statistical analysis and the independent sample t-test, presented in Table 1 and Table 3.

**Table 1.** Statistical description of self and peer assessment scores

| Descriptives | Group | N | Mean | SD | SE |
|---|---|---|---|---|---|
| AME-Value | Peer | 114 | 77.22 | 10.46 | 0.98 |
| | Self | 114 | 78.45 | 9.68 | 0.91 |
| AFE-Value | Peer | 114 | 86.53 | 7.83 | 0.73 |
| | Self | 114 | 86.45 | 8.34 | 0.78 |

Note:

AME is an abbreviation for the mid-semester assessment

AFE is an abbreviation for the assessment of the final semester exam

According to Table 1, the average peer assessment in the midterm exam is lower than the average self-assessment. This did not occur in the final semester exam, where the average assessment between self and peers was nearly identical, namely 86.45 and 86.53. This demonstrates that both students and peers conduct objective assessments in their opinion. The standard deviation shows that the midterm exam assessment by peers is more spread out than the self-assessment. In the final semester examination assessment, self-assessment is more widely distributed than assessment by peers. In addition, there was an increase in the average score of the assessment and a decrease in the standard deviation in the mid-semester and end-semester examinations, both self-assessment and peer-assessment. The decrease in standard deviation in the final semester exam indicates that the distribution of assessment scores is not as broad. This is because of students' awareness or understanding of the assessment criteria provided in the questions.

A prerequisite test, the data normality test, was performed before the independent sample t-test. The results of which are shown in Table 2.

**Table 2.** Data normality test (Shapiro-Wilk)

| | | **W** | **p** |
|---|---|---|---|
| AME-Value | Peer | 0.83 | $< .001$ |
| | Student | 0.84 | $< .001$ |
| AFE-Value | Peer | 0.82 | $< .001$ |
| | Student | 0.79 | $< .001$ |

Note. Significant results suggest a deviation from normality.

According to Table 2, neither the midterm and final semester examinations violate the independent sample t-test normality requirements. In addition, the independent sample t-test was performed, the results of which are shown in Table 3.

16

**Table 3.** Independent Samples T-Test

|  | t | df | p |
|---|---|---|---|
| AME-Value | -0.92 | 226 | 0.36 |
| AFE-Value | 0.07 | 226 | 0.94 |

Note. Student's t-test.

Table 3 shows that for AME, t(226) = - 0.92 and p = 0.36 > 0.001, and for AFE, t(226) = 0.07 and p = 0.94 > 0.001. This result means no statistically significant difference between assessments performed by students and their peers, both in midterm and final semester assessments. In other words, students have the same perception of their contribution to the team. This means that there is a similarity between students and their peers in evaluating the quality of their work and their learning, assessing the extent to which they reflect explicitly stated goals or criteria. These results align with research conducted by Double et al. (2020) and Golightly (2021). This result is different from Kilic (2016), which states a significant difference between the assessment by itself and its peers. In addition, in the midterm assessment, the self-assessment score was higher than the peers'. This result is in line with the results of research, which states that self-assessment scores tend to be higher than peer assessment (Khonbi & Sadeghi, 2013; Golightly, 2021). However, for the final semester examination assessment, the assessment scores by self and peers are relatively the same. This result is different from previous research by Khonbi and Sadeghi (2012) and Chang et al. (2012), who found that self-assessment contrasted with peer-assessment, peer ratings were found to be stronger, with higher scores than self-assessments. This result can be interpreted by making repeated assessments insufficient time. Self and peer assessments will produce relatively the same score. These findings suggest that students undertake well as peer evaluators and that peer assessment is a process with accuracy and consistency (Iglesias Pérez et al., 2020). According to the findings of this and previous studies, there are still differences between self and peer assessment. Some argue that self-assessment is more powerful than peer assessment.

The final one claimed no difference between self-evaluation and peer evaluation. On the other hand, peer assessment is said to be more powerful than self-assessment. This is influenced by the teacher's instructions on how students assess.

**The Relationship Between Self and Peer Assessment and Its Effect on Learning Achievement**

A Pearson correlation analysis is performed to determine whether there is a relationship between self and peer assessments and how it relates to learning achievement. Table 4 summarizes the findings.

**Table 4.** Pearson's Correlations

| Variable |  | AME-S | AME-P | AFE-S | AFE-P | LA |
|---|---|---|---|---|---|---|
| 1. AME-S | Pearson's r | — |  |  |  |  |
|  | p-value | — |  |  |  |  |
| 2. AME-P | Pearson's r | 0.61 | — |  |  |  |
|  | p-value | < .001 | — |  |  |  |
| 3. AFE-S | Pearson's r | 0.28 | 0.19 | — |  |  |
|  | p-value | 0.00 | 0.04 | — |  |  |
| 4. AFE-P | Pearson's r | 0.08 | 0.20 | 0.71 | — |  |

| Variable | | AME-S | AME-P | AFE-S | AFE-P | LA |
|---|---|---|---|---|---|---|
| | p-value | 0.38 | 0.03 | < .001 | — | |
| 5. LA | Pearson's r | 0.35 | 0.33 | 0.41 | 0.45 | — |
| | p-value | < .001 | < .001 | < .001 | < .001 | — |

Note:

AME-S is an abbreviation for mid-semester exam self-assessment.

AME-P is an abbreviation for the mid-semester peer assessment exam.

AFE-S is an abbreviation for end-of-semester self-assessment.

AFE-P is an abbreviation for end-of-semester peer assessment.

LA is an abbreviation for learning achievement.

According to Table 4, there is a high level positive and significant correlation between midterm and final semester assessments by self and peers (r = 0.61; r = 0.71, p < 0.001). This means a decrease or increase will follow a decrease or increase in self-assessment scores in peer assessment scores. These results are consistent with research findings that show a positive and strong correlation between self-assessment and peers (Alias et al., 2015; Nawas, 2020; Zakian et al., 2012). Peer assessment of midterm exams can explain 37. 21% ($R^2$=0.3721) of the variance of student midterm examination assessments. While the assessment of the final semester examination by students themselves can explain 50. 41% ($R^2$=0.5041) of the variance of assessing a peer's final semester exam. This shows that the assessments made by self and colleagues are consistently reliable.

Furthermore, there is a positive and significant correlation with a moderate level between the assessment of the midterm exam by self and peer on student achievement (r = 0.35; r = 0.33, p < 0.001) and between the assessment of the final semester exam by the students themselves and their peers on student learning outcomes (r = 0.41; r = 0.45, p <0.001). This result shows that the assessment by students, self or peer assessment, influences their learning achievement. In other words, the self and peer assessment results are effective in determining the final grade or student achievement. This result aligns with the previous research, which states that self and peer assessments influence learning achievement (Al-Abdullatif, 2020; Birjandi & Tamjid, 2012; Iglesias Pérez et al., 2020; Iraji et al., 2016; Kilic, 2016; Oren, 2018). Student learning achievement can be explained by self-assessment of 12.25% ($R^2$=0,1225) by midterm examination assessments and 16.81% ($R^2$=0,1681) by final semester assessments. These results follow the research results, which states that self-assessment has a positive effect on learning achievement (Sharma et al., 2016; Vasileiadou & Karadimitriou, 2021) and helps get more people involved in your learning activities (Hodgson et al., 2014). Furthermore, peer assessments can explain learning achievement of 10.89% ($R^2$=0,1089) by middle test assessments and 20.25% ($R^2$=0,2025) by final test assessments. This result is under previous research, which states that peer ratings positively influence learning achievement (Brignell et al., 2019; Chang et al., 2020; Chen et al., 2020; Double et al., 2020; Omar. 2019), so that the lecturer can consider the assessment by students in giving final grades of courses with a weight of 10% to 20%.

Students also believe that self-assessment can help them enhance their understanding and passion for the subject, and they find it going to motivate them to gain some self-skills. Academic achievement and learning can be improved because students are intrigued to evaluate their

achievement through self-assessment and are thus motivated to go over the topic again. Self-assessment is an essential component of lessons because self-identification of learners' progress encourages students to progress in learning (Sharma et al., 2016). These results imply that the role of self and peer-assessment positively affects learning outcomes. It shows that students who use self and peer-assessment impact their learning outcomes (Meihami & Esfandiari, 2020). The findings show that self and peer-assessment can be used in mathematics classrooms to help students evaluate their academic achievement. The findings show that both self-assessment and peer-assessment approaches are viable options for evaluating quality student education products (Al-Abdullatif, 2020).

## CONCLUSIONS

Assessment is an essential component of the learning process. Therefore, we need innovation in the assessment process, including the assessment by learners, either by self or their peer assessment. Students' participation in the assessment process will encourage them to be more active and improve their learning quality. According to the findings, there was no significant difference between self and peer assessments in assessing midterm and final exams given by lecturers. However, there is a strong positive relationship between self and peer assessments when evaluating midterm and end-of-semester exams given by the lecturer. The correlation between self and peer assessment and student achievement is moderately positive. Self-assessment and peer assessment contribute to learning achievement. The results of this study imply that the assessments made by students are consistent and can be used as a balance for lecturers in providing final grades for students as their learning achievements.

This study was limited to student and peer assessments of the answers to the lecturers' questions and their relationship with learning achievement. Future research could examine the relationship between student assessments (self and peers) and lecturers and a mathematical model (regression equation) based on self-assessments and peer assessments in predicting learning achievement.

## REFERENCES

Abrache, M.-A., Bendou, A., & Cherkaoui, C. (2021). Clustering and combinatorial optimization based approach for learner matching in the context of peer assessment. *Journal of Educational Computing Research, 59*(6), 1135–1168.

Al-Abdullatif, A. M. (2020). Alternative assessment approaches and quality product design within web-based learning environments. *International Journal of Web-Based Learning and Teaching Technologies*, *15*(3), 60–74.

Alias, M., Masek, A., & Salleh, H. H. M. (2015). Self, peer and teacher assessments in problem based learning: Are they in agreements?. *Procedia - Social and Behavioral Sciences*, *204*, 309–317.

Birjandi, P., & Tamjid, T. N. (2012). The role of self-, peer and teacher assessment in promoting Iranian EFL learners' writing performance. *Assessment and Evaluation in Higher Education*, *37*(5), 513–533.

Brignell, C., Wicks, T., Tomas, C., & Halls, J. (2019). The impact of peer assessment on mathematics students' understanding of marking criteria and their ability to self-regulate learning. *MSOR Connections*, *18*(1), 46–55.

Burkhardt, H., & Swan, M. (2012). Designing assessment of performance in mathematics. *Educational Designer*, *2*(5), 1–40.

Chang, C. C., Tseng, K. H., & Lou, S. J. (2012). A comparative analysis of the consistency and

difference among teacher-assessment, student self-assessment and peer-assessment in a web-based portfolio assessment environment for high school students. *Computers and Education*, *58*(1), 303–320.

Chang, S.-C., Hsu, T.-C., & Jong, M. S.-Y. (2020). Integration of the peer assessment approach with a virtual reality design system for learning earth science. *Computers and Education*, *146*, 103758, 1–32.

Chen, I.-C., Hwang, G.-J., Lai, C.-L., & Wang, W.-C. (2020). From design to reflection: Effects of peer-scoring and comments on students' behavioral patterns and learning outcomes in musical theater performance. *Computers and Education*, *150*, 103856, 1–21.

Double, K. S., McGrane, J. A., & Hopfenbeck, T. N. (2020). The impact of peer assessment on academic performance: A meta-analysis of control group studies. *Educational Psychology Review*, *32*(2), 481–509.

Golightly, A. (2021). Self and peer assessment of preservice geography teachers' contribution in problem-based learning activities in geography education. *International Research in Geographical and Environmental Education*, *30*(1), 75–90.

Goss-Sampson., M. A. (2020). Statistical analysis in Jasp: A guide for students. *4th edition Jasp v 0.14.*

Hassell, D., & Lee, K. Y. (2019). Evaluation of self and peer assessments in a second year engineering module. *IAFOR Journal of Education*, *7*(2), 107–130.

Hodgson, P., Chan, K., & Liu, J. (2014). Outcomes of synergetic peer assessment: First-year experience. *Assessment and Evaluation in Higher Education*, *39*(2), 168–178.

Iglesias Pérez, M. C., Vidal-Puga, J., & Pino Juste, M. R. (2020). The role of self and peer assessment in higher education. *Studies in Higher Education, 47*(3), 1–10.

Iraji, H. R., Enayat, M. J., & Momeni, M. (2016). The effects of self- and peer-assessment on Iranian EFL learners' argumentative writing performance. *Theory and Practice in Language Studies*, *6*(4), 716–722.

Jones, I., & Alcock, L. (2014). Peer assessment without assessment criteria. *Studies in Higher Education*, *39*(10), 1774–1787.

Khonbi, Z. A., & Sadeghi, K. (2012). The effect of assessment type (self vs. peer vs. teacher) on Iranian University EFL students' course achievement. *Language Testing in Asia*, *2*(4), 47–74.

Khonbi, Z. A., & Sadeghi, K. (2013). Self-, peer-, and teacher-assessment: An investigation into Iranian EFL students' attitudes. *Studies in Second Language Learning and Teaching*, *3*(1), 87–107.

Kilic, D. (2016). An examination of using self-, peer-, and teacher-assessment in higher education: A case study in teacher education. *Higher Education Studies*, *6*(1), 136–144.

Lindblom-Ylänne, S., Pihlajamäki, H., & Kotkas, T. (2006). Self-, peer- and teacher-assessment of student essays. *Active Learning in Higher Education*, *7*(1), 51–62.

Meihami, H., & Esfandiari, R. (2020). Comparative effects of self-assessment, peer-assessment, and teacher assessment on efl learners' writing performance. *XLinguae*, *13*(4), 91–107.

Nawas, A. (2020). Grading anxiety with self and peer-assessment: A mixed-method study in an indonesian efl context. *Issues in Educational Research*, *30*(1), 224–244.

Ndoye, A. (2017). Peer/self assessment and student learning. *International Journal of Teaching and Learning in Higher Education*, *29*(2), 255–269.

Omar, S. N. P. (2018). The use of peer assessment to improve students' learning of geometry. *European Journal of Social Science Education and Research*, *5*(2), 187-206.

Oren, F. S. (2018). Self, peer and teacher assessments: What is the level of relationship between them?. *European Journal of Education Studies*, *4*(7), 1–19.

Sharma, R., Jain, A., Gupta, N., Garg, S., Batta, M., & Dhir, S. (2016). Impact of self-assessment by students on their learning. *International Journal of Applied and Basic Medical Research*,

*6*(3), 226-229.

Stančić, M. (2020). Peer assessment as a learning and self-assessment tool: A look inside the black box. *Assessment & Evaluation in Higher Education*, *46*(6), 852-864.

Vasileiadou, D., & Karadimitriou, K. (2021). Examining the impact of self-assessment with the use of rubrics on primary school students' performance. *International Journal of Educational Research Open*, *2*(2021), 1–9.

Zakian, M., Moradan, A., & Naghibi, S. E. (2012). The relationship between self-, peer-, and teacher-assessments of EFL learners' speaking. *World J Arts, Languages, and Social Sciences*, *1*(1), 1–5.