Ed Process Int J

www.edupij.com

ISSN 2147-0901 (Print), 2564-8020 (Online)

EDUCATIONAL PROCESS
INTERNATIONAL JOURNAL

Research Article

Cite this article: Biton, Y. (2025). Student Reflections on Peer Assessments: Benefits and Challenges in a Mathematics Class. *Educational Process: International Journal, 14*, e2025003. https://doi.org/10.22521/edupij.2025.14.3

Received December 01, 2024
Accepted December 31, 2024
Published Online January 02, 2025

Keywords:
Peer assessment, collaborative learning, collaborative evaluation, mathematics education, student attitudes, formative assessment

Author for correspondence:
Yaniv Biton
YanivB@cet.ac.il
Shaanan Academic College, Center for Educational Technology, Tel Aviv, ISRAEL

Student Reflections on Peer Assessments: Benefits and Challenges in a Mathematics Class

Yaniv Biton

Abstract

Background/purpose. This study addresses the challenge of engaging students in meaningful assessment processes within mathematics education, particularly in pre-university preparatory courses. Peer assessment, a growing pedagogical practice, can enhance learning outcomes, promote critical thinking, and improve collaborative skills. However, concerns over fairness, reliability, and students' confidence in their evaluative capabilities often hinder its adoption. This research was conducted to explore students' attitudes toward peer assessment, assess its perceived benefits and challenges, and evaluate its viability as a formative assessment tool in a structured educational context. The main purpose is understanding how peer assessment can contribute to learning while addressing the barriers to its broader implementation.

Materials/methods. The research employed a mixed-methods approach, combining quantitative data from a 29-item attitude questionnaire distributed to 50 students with qualitative insights gathered through semi-structured interviews with 11 participants. Peer assessment activities were conducted in six 90-minute sessions involving problem-solving, developing assessment criteria, and evaluating peers' solutions. Data analysis included statistical evaluations of questionnaire responses and thematic coding of interview transcripts.

Results. Students reported that peer assessment enhanced their learning, improved their understanding of mathematical content and pedagogical requirements, and fostered meaningful peer interactions. However, they expressed concerns about fairness, reliability, and the social dynamics of open peer evaluations. While students valued the educational benefits of the process, they hesitated to accept peer assessment as part of formal grading due to perceived subjectivity and lack of expertise.

Conclusion. While peer assessment is a powerful tool for fostering collaborative learning and critical thinking, its integration into formal grading systems requires careful planning. Addressing training, bias, and anonymity challenges can enhance its reliability and acceptance among students, making it a more robust and equitable practice in education.

1. Introduction

The involvement of students in assessment processes in educational settings has gained increasing prominence in recent years. Peer assessment is a strategy that encourages students to evaluate each other's work, fostering active participation and reflective learning. This method can be applied to evaluate specific task outcomes, individual contributions, or collaborative group efforts (de-Armas-González et al., 2023; Kollar & Fischer, 2010; Li et al., 2019). Peer assessment may serve formative or summative purposes. Formative peer assessment allows for early identification of errors and misconceptions, facilitating deeper understanding and knowledge retention. Typically, the criteria for assessment are predefined and align with the formal requirements specified for the evaluated task.

The ongoing development of peer assessment as an instructional tool requires continuous exploration of how students' attitudes evolve over time and what factors contribute to their acceptance or resistance. Researchers such as Van Zundert et al. (2010) have called for further investigation into the long-term effects of peer assessment on learning, as well as clearer methodological frameworks that link the assessment process to specific learning outcomes. These efforts aim to solidify the role of peer assessment as a meaningful, reliable, and equitable practice in modern education. Considering these calls, this study aimed to investigate the attitudes of students enrolled in a preparatory course in mathematics at the Center for Pre-University Education of the Technion – Israel Institute of Technology toward peer assessment using data gathered through peer assessment activities where students engaged in solving problems, creating evaluation criteria, and assessing their peers' solutions.

2. Literature Review

Peer assessment has been defined in multiple ways throughout the literature. For instance, Falchikov (1995) characterizes it as a process where individuals or groups rate their peers, while Topping (1998) offers a more detailed definition, describing it 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" (p. 250). These definitions highlight the evaluative nature of peer assessment, positioning students as both assessors and learners in a collaborative framework.

2.1 Benefits and Challenges of Peer Assessment

Numerous studies have demonstrated that peer assessment is an intellectually demanding task that enhances learning and boosts motivation (Bedford & Legg, 2007; Black & Wiliam, 2006; Orsmond et al., 2000). As an alternative assessment tool, it fosters interactions between students and between students and instructors, thereby increasing learners' understanding of the material through the perspectives of their peers. Additionally, it has been implemented across a variety of disciplines, including computer science (Venables & Summit, 2003), mathematics-teacher training (Zevenbergen, 2001), psychology (Smith et al., 2002), and economics (Brindley & Scoffield, 1998). The application of peer assessment in these diverse fields highlights its versatility and its potential to improve learning outcomes across educational contexts.

Peer assessment supports cognitive and metacognitive development by providing students with opportunities to give and receive feedback. The process of offering feedback requires the assessor to thoroughly understand the objectives of the task, the criteria for success, and how to relate the outcomes to these standards (Xiang et al., 2021; Yin et al., 2022). As noted by Zariski (1996), the assessment process is as significant as its results, enabling learners to engage in critical thinking and reflective practices. For example, Kollar and Fischer (2010) emphasized that peer assessment allows students to gain insights into the thinking processes of their peers, thus fostering a greater sense of

https://doi.org/10.22521/edupij.2025.14.3 Published online by Universitepark Press

responsibility for their learning. This dynamic encourages students to view the task from a teacher's perspective, which aids in grasping the intricacies of the evaluation process.

However, it is crucial to note that the integration of peer assessment is not without its challenges. Students may experience anxiety over potential exposure, fear of offending, or lack of trust in the fairness of the process (de-Armas-González et al., 2023; Kirkpatrick & Fuller, 1995). To mitigate these concerns, it is essential to provide adequate training and clear guidelines, as studies suggest that experience and familiarity with peer assessment can lead to more positive attitudes towards its implementation (Cheng & Warren, 1997; Wen & Tsai, 2006).

Recent meta-analytical research has highlighted the nuanced impact of peer assessment on student learning across various educational contexts. According to Li et al. (2019), peer assessment generally significantly improves student performance, with an effect size of 0.291 standard deviation units over those who did not participate. Key factors influencing the effectiveness of peer assessment include the presence of rater training and the use of computer-mediated assessment methods. When students received explicit training on how to assess, the effect size was markedly larger, suggesting that structured training enhances both the accuracy and educational benefit of peer assessment. Additionally, peer assessments conducted digitally were found to be more effective than traditional, paper-based methods, underscoring the role of technology in facilitating clearer, more structured feedback. This evidence underscores the importance of designing peer assessment activities incorporating structured training and leveraging digital tools to maximize learning outcomes (Li et al., 2019).

Moreover, recent studies have further explored the mechanisms of peer assessment, particularly within educational contexts that focus on developing critical problem-solving skills. For instance, de-Armas-González et al. (2023) investigated how pre-service secondary mathematics teachers engaged in assessing their peers' solutions to open-ended mathematical problems. Their study revealed that the assessment process is often influenced by the assessors' own problem-solving approaches, which can challenge the objectivity of evaluations. Nevertheless, the research showed that peer assessment activities help future teachers enhance their mathematical understanding by exposing them to diverse problem-solving strategies. This underscores the value of incorporating peer assessment practices into teacher-training programs to build comprehensive evaluation skills and facilitate deeper learning (de-Armas-González et al., 2023).

2.2. Student Attitudes Toward Peer Assessment

Research has extensively examined student attitudes toward peer assessment, particularly concerning how their experiences and familiarity with the process influence their perceptions. Many studies indicate that students generally develop favorable attitudes toward peer assessment, especially after having participated in it (Cheng & Warren, 1997; Strachan & Wilcox, 1996; Venables & Summit, 2003; Wen & Tsai, 2006). For instance, Cheng and Warren (1997) found that training and experience significantly improved students' comfort levels and perceptions of fairness when assessing their peers. Their study of first-year electrical engineering students demonstrated that students felt more confident and comfortable with peer assessment when it followed structured training and practical engagement and viewed it as a valid component of their learning.

A study by Wen and Tsai (2006) investigated university students' attitudes toward peer assessment with both traditional and online formats. The researchers developed a 20-item questionnaire based on four subscales: Positive Attitudes, Online Attitudes, Understanding-and-Action, and Negative Attitudes. The study's findings revealed that students generally hold positive views about peer assessment, appreciating it as a tool to enhance interaction and understanding between peers and instructors. Male students exhibited more positive attitudes than female students, and those with prior experience in peer assessment showed fewer negative perceptions.

https://doi.org/10.22521/edupij.2025.14.3 Published online by Universitepark Press

Moreover, while students recognized the utility of online peer assessment for logistical purposes, they did not necessarily see it as a beneficial learning tool. The study highlights the need for clear guidelines and effective implementation strategies to ensure the reliability and validity of peer assessment processes. In the present study, we rely on Wen and Tsai's questionnaire as a foundational tool to assess attitudes toward peer assessment.

The literature also underscores the importance of context in shaping students' attitudes. For example, in a study involving computer science undergraduates, Venables and Summit (2003) found that students who initially expressed reservations about peer assessment became more positive after experiencing it firsthand. Through anonymous evaluations of literature reviews, students learned to appreciate the diverse perspectives of their peers and recognized the value of engaging with varying levels of analysis and critique. This shift in attitude aligns with the principles of reflective learning, where the act of evaluating others deepens one's own understanding of the subject matter.

3. Methodology

The primary objective of this study was to explore and understand students' attitudes toward peer assessments in mathematics education. To address this, the research focused on the question: "What are students' attitudes toward peer assessments that they perform in mathematics education?" The study employed a mixed-methods approach, combining qualitative and quantitative data collection techniques. We conducted 11 in-depth interviews with a sampling of students, taking detailed field notes during these sessions and engaging in informal conversations to gain deeper insights. Additionally, an attitude questionnaire, adapted from the study by Wen and Tsai (2006), was distributed to 50 students who participated in peer assessment activities. Since the current study focused exclusively on classroom-based assessments, our questionnaire omitted the Online Attitude Subscale used in Wen and Tsai's original research. The analysis in this study aimed to provide a comprehensive view by examining the reliability of responses across three different scales, categorizing the distribution of attitudes into positive, neutral, or negative, and using qualitative findings from the interviews to contextualize and support the quantitative data. These methods allowed for a thorough exploration of student perspectives on the practice of peer assessment in a structured educational setting.

3.1. Research Context and Participants

The study took place at the Center for Pre-University Education of the Technion – Israel Institute of Technology. The participants comprised approximately 30 students in each of the two classes that taught an intensive, single-semester preparatory course in mathematics to bring the students up to level five-matriculation proficiency. Such preparatory courses are offered in many universities worldwide to close any gaps between students' achievements in high school and prerequisites for academic study. Typically, the students' progress is assessed systematically using traditional methods, i.e., exams, but during this course, students worked in groups and were asked to assess their peers' work. One teacher taught both classes. Six peer assessment activities (three from each class) were recorded using a video camera and audio recordings. The study took place in Hebrew and transcripts were translated for this study.

3.2. Peer Assessment in the Current Study

Each of the six peer assessment activities was 90 minutes long. In the first stage (25 minutes), the students were asked to individually solve a mathematical problem randomly chosen from the course textbook (see Figure 1 for a sample problem).

https://doi.org/10.22521/edupij.2025.14.3 Published online by Universitepark Press

$f(x) = \frac{\cos x}{\sin^2 x}$ is a function at $-\frac{\pi}{2} \leq x < 0$.

a) Find all the asymptotes, if any, which are parallel to the axes.

b) Find the points of intersection, if any, of the function with the axes.

c) Prove that the function is increasing.

d) Find the equation of the tangent to the graph of f(x) at its point of intersection with x-axis.

e) Find the area of the figure constrained by f(x), the above tangent, and x = –0.5.

Figure 1. Sample Problem

In the second stage (15 minutes), the students were randomly divided into groups of two or three. They were asked to formulate criteria for evaluating the other students' solutions to the problem and to weigh the criteria according to their relative importance. Each group was given a four-page booklet in which they were to write their evaluation criteria on the first page and use the remaining pages to comment and score the other students' solutions.

In the third stage (50 minutes), each group received the solutions of three other students for assessment. This was not anonymous (i.e., all the students knew whose solutions they were reviewing and who was checking theirs). The inter-group conversation was allowed and encouraged at all stages of the activity. The students were advised that this was merely an exercise and that their assessment would not affect their peers' final marks. This was to prevent them from feeling inhibited about giving a bad mark or comment. After evaluation, the annotated work was returned to the assessees, who were given the opportunity to respond if they wished to explain their work or question the assessment.

3.3. Designing Peer Assessment Activities for this Study

The way the students designed their peer assessment activities was affected by the way they engaged in social interaction, took responsibility, and depended on each other. The nature of the activity provided the learners with learning situations in which they were given roles that required them to provide explanations and ask questions alongside their role as learners, that is, the need to learn about the correct solution to the problem and to assess it appropriately. These settings evoked discussion and active listening throughout the phases of the assignment. The students in each group were chosen randomly, and it appears that this allowed them to learn from, help, and enrich each other. The environment created was conducive to active involvement in the assessment assignment. Each student was able to show their skills and contribute knowledge to the group. This social interaction in collaborative learning, where a student is exposed to multiple perspectives and exchange of knowledge among the members of the group, is known to have a positive impact on learners' achievements (Kollar & Fischer, 2010; Leikin & Zaslavsky, 1999; Radford, 2011).

3.4. Challenges in Integrating Peer Assessment Activities

Alongside the advantages of collaborative learning for the learner, there are also challenges involved in integrating peer assessment into the curriculum. However, these challenges may be overcome by designing peer assessment activities as in the current study.

3.4.1. First challenge: Group pressure

Students need to feel that they are members of the team when working in a collaborative group. Otherwise, they may feel pressured by the group and, as a result, cannot distance themselves and express themselves through a personal product that characterizes them.

https://doi.org/10.22521/edupij.2025.14.3 Published online by Universitepark Press

As stated above, the members of the collaborative groups were first told to solve the assignment on their own. This stage may be considered a kind of "test writing," and it contributed to the discourse between the assessors since each brought their own position concerning their own individual way of solving the problem. They may have encountered difficulties or observed difficulties encountered by others. However, this stage enabled brainstorming, which proved valuable at the stage where the criteria for assessing the solutions were set, as well as when they discussed the correctness of the solutions they assessed.

3.4.2. Second challenge: Advanced students

A small collaborative group reduces the students' average to that of the group. The need for conformity in the group mainly harms the advanced student and limits their thinking skills and creativity.

Similar to other studies on collaborative learning that point to group heterogeneity as an effective tool for enhancing collaborative learning (Davidson, 1990; Johnson & Johnson, 1985), we observed that when groups were composed of students who succeeded in completing the assignment and those who could not, the heterogeneity paid off. The successful students (including those who only partly succeeded) very quickly adopted the role of "teacher" within the group. They started to answer their peers' questions concerning the solutions they assessed. This situation, where one group member explains the material to their peers, is well known and supported in the literature. It is also an advantage for the "teacher" because the act of explaining may shed light on aspects of which they were previously unaware, and they can also better understand what is unclear to their peers. All the group members speak a common language, thus the person receiving the explanation gains better understanding (Webb, 1991).

3.4.3. Third challenge: Less motivated students

In a collaborative setting, weak, lazy, or uninterested students benefit from collaborating with the more diligent ones, thus relinquishing their responsibility for their own learning.

Various options for conducting the peer assessment tasks were tried as part of the exploratory study. One was a situation where a single student solved the problem, and another assessed it. Here, we noticed that students who failed to solve the problem on their own refrained from participating in the peer assessment, claiming that they could not assess a solution because they had failed to find themselves. The phenomenon of social avoidance during collaborative learning is well-known in the literature (Kerr & Tindale, 2004). However, the present study did not observe this phenomenon. In fact, we noticed that students who had difficulty solving the assignment in the initial, individual phase still took an active part in the self-assessment phase, e.g., asking the group members questions, making attempts to probe deeper and to understand the phases of the solution offered; and then, later, to take a stand on its validity, even voicing their opinions concerning the grade they had to give.

In this study, peer assessment was done openly. In such cases, fear of exposure and personal hurt, fear of hurting one's peers, and difficulty being objective are well-known phenomena recognized in the literature. However, in this study, the grades resulting from the peer assessment did not influence the formal ones, thus reducing these fears for the students, as will be seen in the study's findings. Note that the participants initially expressed concerns about their assessments influencing their peers' final grades. However, once they were assured this would not happen, they initiated assessor-assessee meetings spontaneously.

3.5. Research Tools and Data Analysis

The study employed a mixed-methods approach (Creswell, 2003). For the quantitative component, a 29-item questionnaire based on the work of Wen and Tsai (2006) was utilized. The

https://doi.org/10.22521/edupij.2025.14.3 Published online by Universitepark Press

amended questionnaire measured attitudes across three subscales: Positive Attitudes, Understanding and Action, and Negative Attitudes. Responses were rated on a Likert scale ranging from 1 = "Strongly Disagree" to 5 = "Strongly Agree." (Note: several statements were negatively worded, meaning a lower score reflected a more positive attitude.) The questionnaire was translated into Hebrew, and the accuracy of the translation was verified through a rigorous process involving two expert educators fluent in both Hebrew and English, who independently reviewed the translation to ensure that it maintained its original meaning, clarity, and intent and that it was linguistically accurate and culturally appropriate for the target audience. Discrepancies, if any, were discussed and resolved to achieve consensus.

For the qualitative component, semi-structured interviews were conducted with 11 students after they had participated in the peer assessment tasks. The students were randomly selected. The interviews aimed to explore their attitudes toward the value, advantages, and disadvantages of peer assessment; the challenges they encountered during the evaluation; what potential learning benefits might be associated with the process; their feelings regarding evaluating their peers; whether they believed that their assessments were objective enough to be included in their classmate's final mark; and so on. We also informally interviewed the teacher to obtain her reactions to the statements.

All the interviews were fully recorded and transcribed, then analyzed using an inductive approach. Key categories were derived from the transcripts, reflecting the core themes from the students' responses. Both the questionnaire and interviews provided a comprehensive view of the students' experiences with peer assessment, offering an in-depth understanding of their perspectives considering the benefits and challenges they associated with this process.

4. Results

This section outlines the key findings, highlighting the benefits, challenges, and overall perceptions observed during the peer assessment activities. The data provide a detailed understanding of how students engage in peer assessment and the factors influencing the assessment's effectiveness in a learning environment. In the first section, we present the reliability results of each of the three subscales and compare them to those of Wen and Tsai (2006).

4.1. Reliability Results of the Three Subscales

4.1.1. Positive attitude subscale

The Positive Attitude Subscale (Table 1) comprises eight statements (Q2–4, 10, 11, 13, 14, 24). The value of α (Cronbach alpha) is 0.70. Thus, the statements can be considered to have internal consistency; as seen in Table 1, removing any single question would not increase the value of α. Therefore, the eight statements can be used to define one Positive Attitude Subscale.

https://doi.org/10.22521/edupij.2025.14.3 Published online by Universitepark Press

Table 1. Analysis of the Positive Attitude Subscale Statements

Cronbach's alpha

Statement number	Statement text	Raw variables		Standardized variables	
		Correlation with total	Alpha	Correlation with total	Alpha
Q2	Peer assessment is beneficial to my learning.	0.33	0.69	0.34	0.70
Q3	Peer assessment helps me better understand the teacher's requirements.	0.36	0.69	0.35	0.69
Q4	Peer assessment activities can help improve my skills in verbal communication.	0.37	0.69	0.37	0.69
Q10	Peer assessment activities motivate me to learn.	0.45	0.66	0.44	0.68
Q11	Peer assessment activities increase the interaction between me and my teacher.	0.38	0.68	0.39	0.68
Q13	Peer assessment activities help me develop a sense of participation.	0.58	0.64	0.59	0.64
Q14	Peer assessment activities increase the interaction between me and my classmates.	0.42	0.67	0.41	0.68
Q24	I think using peer assessment to assess students' performance is fair.	0.32	0.69	0.34	0.70

4.1.2. Understanding and action subscale

The Understanding and Action Subscale (Table 2) measures students' understanding of others' ideas and what they should do during the activity. This subscale comprises three statements (Q5, 22, 25). The α is very low (0.14). Table 2 shows that one of the items interferes with the content. Therefore, it is impossible to formulate one Understanding and Action Subscale variable from the

https://doi.org/10.22521/edupij.2025.14.3 Published online by Universitepark Press

three statements in their present condition. Omitting item Q22 increases α to 0.51, thus indicating medium-low internal consistency between statements Q5 and Q25.

Table 2. Analysis of the Understanding and Action Subscale Statements

Cronbach's alpha					
Statement number		Raw variables		Standardized variables	
	Statement text	Correlation with total	Alpha	Correlation with total	Alpha
Q5	Peer assessment activities help me understand what other classmates think.	0.13	-.05	0.16	-.05
Q22	The teacher should develop the criteria for peer assessment activities for the students.	-.10	0.51	-.10	0.54
Q25	Students should participate in the development of criteria for peer assessment activities.	0.23	-.33	0.26	-.33

4.1.3. Negative Attitudes Subscale

The Negative Attitudes Subscale (Table 3) comprises four statements (Q18, 27–29). The value of α is 0.54. However, the table shows that the value of α cannot be increased by removing any one of the statements. Thus, using the four-statement Negative Attitudes Subscale variable for negative attitudes toward peer evaluation may be slightly problematic.

The findings show that the reliabilities of both the Understanding and Action Subscale and the Negative Attitudes Subscale are low. However, a high internal consistency should not be expected since some of the statements in these subscales express attitudes, and others report on performance. Therefore, the results for each statement must be checked individually, as is demonstrated in the next section. Also note that in the current study, everyone who responded to the questionnaire had participated in peer evaluation, in contrast to Wen and Tsai's study, where some of the students who responded to the questionnaire had not participated in the peer evaluation. Therefore, their responses to statements that refer to performance are only attitudes.

https://doi.org/10.22521/edupij.2025.14.3 Published online by Universitepark Press

Table 3. Analysis of the Negative Attitudes Subscale Statements

Statement number	Statement text	Raw variables		Standardized variables	
		Correlation with total	Alpha	Correlation with total	Alpha
Q18	I think students should not be responsible for making assessments.	0.35	0.46	0.35	0.46
Q27	Peer assessment is time-consuming.	0.24	0.54	0.23	0.56
Q28	The marks I give to classmates are related to the marks I received.	0.44	0.38	0.46	0.37
Q29	If I receive lower marks than I expected, then I will give lower marks to my classmates.	0.30	0.49	0.30	0.50

4. 2. Distribution of the Scores of the Various Statements into Three Levels: Positive, Neutral, or Negative

To check whether the respondents felt positively regarding peer evaluation, a T-test was conducted to find if the average level of agreement on the various indices was higher than neutral. A neutral position is scored zero; agreement received a positive score (1, 2), and disagreement a negative score (-1,-2). The averages and results of the T-test are displayed in Table 4.

Table 4. Measure of Average Agreement in the Three Subscales

Sub-scale	N	Mean	St. Dev.	Median	t Value	Pr > |t|
Positive Attitude	50	0.21	0.52	0.25	2.85	0.0064
Understanding and action	50	0.52	0.87	0.75	4.14	0.0001
Negative Attitude	50	-0.43	0.60	-.375	-5.10	<.0001

Table 4 shows that the scores differ significantly from zero (neutral) for all three subscales. However, one must know that a negative score for the Negative Attitude subscale implies a tendency toward attitudes that favor peer evaluation. Since the evaluation averages are not high enough to indicate strong positive attitudes, the same test was conducted for each item in the questionnaire. Thus, it is possible to see which attitudes are significantly positive or negative. Table 5 shows the statements that received significantly positive (1 or 2) or negative (-1 or -2) scores (neutral statements were eliminated).

https://doi.org/10.22521/edupij.2025.14.3 Published online by Universitepark Press

Table 5. The measure of Average Agreement in Each Statement

Item no	Item statement	N	Mean	St. dev.	Standardized	t–value	Pr > |t|
Q2	Peer assessment is beneficial to my learning.	50	0.40	0.90	1	3.13	0.0029
Q3	Peer assessment helps me better understand the teacher's requirements.	50	0.84	0.65	1	9.13	<.0001
Q5	Peer assessment activities help me understand what other classmates think.	49	0.51	0.98	1	3.64	0.0007
Q6	Peer assessment activities make me nervous and put me under pressure.	50	-0.88	1.15	-1	-5.39	<.0001
Q7	Peer assessment activities help the teacher to understand what students think.	49	0.44	0.89	1	3.53	0.0009
Q9	I am fair to other classmates during peer assessment activities.	50	0.82	1.00	1	5.78	<.0001
Q14	Peer assessment activities increase the interaction between me and my classmates.	50	0.58	0.83	1	4.91	<.0001
Q15	I think the opinions given by my classmates are acceptable.	50	0.66	0.68	1	6.78	<.0001
Q16	I think I am not being fair enough to my classmates during peer assessment activities.	50	-0.86	1.14	-1	-5.32	<.0001
Q20	Peer assessment activities will affect my relationships and friendships with my classmates.	50	-0.48	1.03	-1	-3.28	0.0019
Q25	Students should participate in the development of criteria for peer assessment activities.	50	0.56	1.01	1	3.91	0.0003
Q26	Peer assessment activities are fair if done anonymously.	50	0.60	1.03	1	4.12	0.0001
Q27	Peer assessment is time-consuming.	50	0.80	0.92	1	6.11	<.0001
Q28	The marks I give to classmates are related to the marks I received.	50	-0.94	0.84	-1	-7.88	<.0001
Q29	If I receive lower marks than I expected, then I will give lower marks to my classmates.	50	-1.48	0.81	-2	-12.80	<.0001

https://doi.org/10.22521/edupij.2025.14.3 Published online by Universitepark Press

4.3. Qualitative Findings

The 11 students were interviewed after the quantitative assessment. The purpose of the interviews was to help further understand the students' attitudes to the peer assessment exercise. Their verbal statements were listed and matched with statements in Table 5 that corresponded. They were then divided into the following six categories (some categories overlap). Sample quotes were added for further illustration.

4.3.1. Category 1 - Enhancing learning outcomes: Understanding the mathematical content and how to solve problems better

Q2, 25. The students reported that participating in the assessment program, especially in determining the assessment criteria, helped them understand the mathematical concepts better. They learned how a question may be divided into components, thus leading them to consider strategies to help solve similar questions because "when you look at a question and consider how the teacher will distribute the points, you can solve according to that."

Q5. They also appreciated how their exposure to their peers' thinking allowed them to realize that there may be multiple ways to solve any particular problem. They pointed out that when assessment is test-based, they receive their grade and have no opportunity to really understand the concept or "observe the work of my peers and compare our work." With peer assessment, however, they can compare different methods, thus expanding their understanding of the concept under study. This occurred even if their peers erred, as by understanding the source of their peers' errors, "I can get something out of that for me... I actually learn about mistakes."

4.3.2. Category 2 - Understanding their teacher's and the institution's criteria

Q3. The students reported that the activity helped them perceive the problem from the teacher's point of view and engage in meta-cognitive reflection. It encouraged them to "look at the exercise from the teacher's side and understand the [didactic] goal." Cultivating such thoughts may help students deal with other problems in the future since although, for a student, the "simple goal is to get a good grade, " by understanding the criteria a teacher is looking for in a "good" answer "you really see what it consists of, how it is structured, and what to emphasize."

Q3. The students pointed out that the collective development of criteria allowed them to understand better the demands of the teacher or the institution and whether "the way I assessed the person … corresponds with what the institution requires." One student noted that he had graded "demonstration of knowledge" for 50% of the mark but discovered the institution only values it at 5 to 40 percent. "I have to ensure that the institution and I are on the same page."

The teacher, too, noted that the students, when discussing criteria, attributed a more significant percentage of the grade to components that seemed less important to her and vice versa. She pointed out, "They evaluated the use of formulae and equations at 50%. I would prefer that they give 50% for the solution plan." She explained that "equations and formulae are tools and not the essence. They are secondary." She also commented that this made her aware that, in the future, she needed to emphasize this to her students.

Q7. This concept also worked in the opposite direction. Some students mentioned that collectively defining criteria was also valuable for the teachers, as it made them more aware of what the students considered to be the problem's components, thus allowing them to coordinate expectations and even understand if they lacked some crucial knowledge. As one student put it, "If I express what I expect to get or what I think is appropriate, I think it will help the teacher see what I see… to coordinate between teacher and student what they expect from the other and themselves."

https://doi.org/10.22521/edupij.2025.14.3 Published online by Universitepark Press

They also expressed their satisfaction that it allowed "the institution an opportunity to observe [the student's attitudes] and respond. It is an opening for the student to speak."

The teacher also noted the advantages that can be found by having to formulate the criteria: "Since they have to allot points to each part of the problem, they first have to organize it and divide it into the parts that must be solved." She noted that this was good practice because during a "real test," students are often under stress and first tackle what they know without dividing the question into parts and organizing the thoughts in their heads. This may mean that they waste valuable time on unimportant parts. "However, by understanding the concept of allotting points, they will see the importance of organizing their work."

Some students expressed that defining criteria is difficult as they are not sure if they are adequately qualified for this task: "The drawback is that we do not know exactly what criteria are important for evaluation." (See also Category 5, Reliability). Q27 indicates that the students felt that the evaluations took up much time. This may lead them to better appreciate the time and effort teachers invest in assessing students' work.

4.3.3. Category 3 - Peer interaction and social conflict

Q5, 15. The students noted that the activities increased (positive) interaction with their classmates. As a result of the activity, they could listen to and understand their classmates' thinking and learn to respect opinions that differed from theirs but were still valid.

Q6, 14, 26. As mentioned, the students were informed in this study that their assessments would not affect their classmates' final grades. This likely helped remove their inhibitions and stress about making an unfair or incorrect assessment and see the activity as educational.

Q9, 14. Because these were "open" evaluations, the students were aware of who had evaluated them and the scores they had assigned. This, they pointed out, could lead to various difficulties. For example, although the students believed that they were, for the most part, fair in their assessments, the "open" evaluations had sometimes led to arguments and objections among the students: "There should be no conflict in a classroom; it does not contribute"; "In one way or another, the credibility of the evaluator comes up for discussion … it becomes a problem."

Q16, 20, 28, 29. They also said that open evaluations could lead to competition, feelings of offense, and a lack of objectivity: "If you ask your peer how they scored you and they answer, say, 65%, you might ask, 'Why so low? In that case, I will also give you [a low grade].'" On the other hand, being in the same social group might have the opposite effect: "If you evaluate someone who is your friend and he solved it incorrectly, you may give him more points than he really deserves"; "… you do not want to give low grades to a friend"; "You may grade based on what you know about the student. If you know, they usually have good grades…you immediately decide that he must know and grade accordingly"; "it is not impartial." They also mentioned that if they had a strained relationship with one of their peers, they did not trust that student to grade them fairly.

Notwithstanding, we noted that some students suggested that anonymous assessment might remove inhibitions, make it easier for them socially, and even increase the reliability of the process: "If it were anonymous, then it would have been easier for us"; "If I were checking someone from a different class, then I would really invest [effort] in checking what grade he should or should not get." If they felt this way, they could be more objective as they would not worry about offending.

4.3.4. Category 4 - Fairness

Six statements (Q6, 9, 16, 20, 28, 29) dealt with the student's sense of fairness in dealing with an open assessment. Q9 received a score of +1. However, the other five were negatively worded statements; therefore, their negative scores should be considered positive. In other words, all six

https://doi.org/10.22521/edupij.2025.14.3 Published online by Universitepark Press

statements suggest that, overall, students did not feel nervous or under pressure (Q6), felt they were being fair (Q9, 16), did not believe that the activity would affect their relationship with their classmates (Q20), or that the mark they received would influence how they would score their peers (Q28, 29). Nevertheless, Q26 indicated that many students thought the assessment would be fairer if done anonymously.

4.3.5. Category 5 - Reliability and confidence

Despite the overall positive scores and the many positive reactions offered during the interviews, some students still expressed concern with the concept of peer assessment. One concern was their inexperience in the discipline, meaning that they felt they might not have enough foundational knowledge to properly understand the objectives of the task, understand the intricacies of the evaluation procedure, and properly evaluate their peers, especially since they, the "evaluators," basically had the same level of knowledge as the "evaluatees": "We do not have enough knowledge to evaluate other students' tests"; "We [really] do not know the absolute final answer … do not have a solid basis on which to evaluate"; " A person who evaluates someone else needs to be an expert in that field"; "We do not know the material to the same extent the teacher knows it."

4.3.6. Category 6 - Hesitation using peer evaluation for formal scores

Q26. An important finding from the interviews was the students' feelings that it was inappropriate to integrate their peer evaluation into the formal assessments. This concern was mainly due to their lack of confidence regarding the reliability of their assessment; the students felt that although it was a good exercise and taught them a lot, they did not feel comfortable having their assessment used as part of the official formative assessment.

5. Discussion

The insights gained from this study contribute to the broader discourse on collaborative learning and offer practical recommendations for improving peer assessment practices in educational settings.

5.1. Benefits of Peer Assessment in Enhancing Learning Outcomes

The findings clearly demonstrate that peer assessment can be a valuable tool for enhancing learning outcomes in a pre-university mathematics course. Students who participated in peer assessment activities reported several key benefits, including increased understanding of the pedagogical requirements, improved interaction with peers, and exposure to a range of problem-solving approaches. These findings align with previous studies that have emphasized the value of collaborative learning environments (de-Armas-González et al., 2023; Kollar & Fischer, 2010; Radford, 2011).

The process of evaluating peers' work forced students to adopt a metacognitive approach to their learning. As they examined different solutions, they were required to think critically about the steps taken by their peers, which often involved revisiting and reflecting on their own understanding of the mathematical concepts. This aligns with the findings of Topping (2017) and Li et al. (2019) that peer assessment encourages students to articulate their thinking, which leads to deeper comprehension. The act of providing feedback also required students to engage in clear and precise communication, which is a crucial skill in the study of mathematics.

Moreover, the data from both the questionnaires and interviews suggest that students perceived the peer assessment process as beneficial for developing their analytical skills. Many students reported that seeing how their peers approached problems helped them understand different strategies and methods, thus broadening their own problem-solving toolkit. This exchange of diverse perspectives is a key advantage of collaborative learning, as it encourages students to consider

https://doi.org/10.22521/edupij.2025.14.3 Published online by Universitepark Press

multiple approaches to a single problem, fostering adaptability and creative thinking (de-Armas-González et al., 2023; Leikin & Zaslavsky, 1999).

5.2. Challenges and Concerns: Fairness, Reliability, and Social Dynamics

Despite the clear benefits, the study also highlighted several challenges associated with implementing peer assessment. The main concerns students raised were fairness, reliability, and the social dynamics involved in evaluating their peers' work. While peer assessment was generally perceived as a positive learning experience, students were hesitant to see it integrated into their formal grading system due to these issues.

One of the primary challenges was students' lack of confidence in their own ability to assess their peers accurately, and some participants expressed concern that they did not possess the requisite knowledge or skills to provide valid evaluations. This issue of perceived competence aligns with previous findings by Webb (1991), which suggest that students may feel unqualified to take on the role of evaluator, especially in subjects where they are still building their own foundational knowledge. The data from the study suggest that this challenge could be mitigated by providing students with more training and clear guidelines on how to assess their peers effectively.

Another significant concern was related to the potential bias in peer assessment. Students feared that evaluations could be influenced by personal relationships, leading to "friendship marking," as described by Kollar and Fischer (2010). Indeed, the qualitative data revealed that some students did feel uncomfortable providing low marks to friends or, conversely, feared receiving harsh evaluations from peers with whom they had strained relationships. This social pressure can compromise the integrity of the peer assessment process; it highlights the need to develop strategies to ensure anonymity and objectivity in peer evaluations.

5.3. Recommendations for Improving Peer Assessment Practices

The findings of this study suggest several ways in which peer assessment practices could be improved to maximize their benefits while minimizing the challenges. First and foremost, the implementation of anonymous peer assessments could address many of the concerns related to fairness and bias. Ensuring that students do not know whose work they are evaluating reduces the risk of social bias, which can lead to more honest and objective feedback. This recommendation is consistent with previous research, which has found that anonymity in peer assessment can enhance the reliability of the evaluations (Sambell & McDowell, 1997).

Another recommendation is the integration of training sessions to prepare students for their role as peer assessors. These sessions could focus on providing constructive feedback, using evaluation criteria effectively, and approaching the assessment process more objectively. It has been shown that training can improve the effectiveness of peer assessment by equipping students with the skills they need to assess their peers' work confidently and accurately (Strijbos & Sluijsmans, 2010).

Finally, when students participate in formulating the criteria, they gain a better understanding of what constitutes a good solution. This can enhance their own learning and help them provide more meaningful feedback to their peers. This collaborative approach to setting evaluation standards can also help align students’ expectations with the instructors', thus reducing discrepancies in grading and fostering a more cohesive learning environment.

5.4. Limitations

While this study provides valuable insight into the role of peer assessment in mathematics education, certain limitations must be considered. First, the study was conducted within the specific context of a pre-university preparatory course at the Technion–Israel Institute of Technology. This specific educational setting, with its particular student demographic, may limit the generalizability of

https://doi.org/10.22521/edupij.2025.14.3 Published online by Universitepark Press

the findings to other contexts, such as secondary schools, university-level courses, or different countries. Additionally, the students were aware that their peer assessment activities would not affect their final grades, which may have influenced their attitudes and behaviors during the assessment process. Future research could explore the impact of integrating peer assessment into formal grading to determine if the findings hold true in a higher-stakes environment.

6. Conclusion and Implications

The study offers convincing confirmation that peer assessment fosters a deeper understanding of mathematical concepts, improves problem-solving skills, and enhances analytical abilities through reflective evaluation, but also involves challenges concerning fairness, reliability, social dynamics, bias, and confidence in evaluating peers. Recommendations include implementing anonymous peer assessments, providing assessor training, and involving students in formulating evaluation criteria. These measures should maximize the benefits of peer assessment while addressing its challenges.

This study makes several important contributions to the field of educational research. It provides empirical evidence of the benefits of peer assessment as a tool for enhancing learning. It highlights how it can foster deeper understanding, improve communication skills, and encourage collaborative problem-solving. The study also sheds light on the challenges associated with peer assessment, particularly issues related to fairness, reliability, and social dynamics. By addressing these challenges, educators can better implement peer assessment practices that are both effective and equitable.

Furthermore, the study emphasizes the importance of involving students in the assessment process, not only as passive recipients of feedback but as active participants who contribute to the learning environment. This approach aligns with modern pedagogical practices that advocate student-centered learning, where students take greater responsibility for their own learning and the learning of their peers. As educational institutions continue to explore ways to integrate formative assessment strategies, this study provides a framework for how peer assessment can be effectively implemented.

Building on the findings of this study, several directions for future research are recommended. First, further studies might examine the long-term effects of peer assessment on students' learning outcomes and their development of metacognitive and analytical skills. For example, longitudinal studies could explore how repeated engagement with peer assessment influences students' ability to critically evaluate mathematical solutions over time.

Second, future research could investigate the integration of peer assessment into formal grading systems and its potential impact on students' perceptions of fairness and motivation. This includes exploring the use of technology to support anonymous peer evaluations and ensuring that social biases do not undermine the assessment process.

Third, exploring peer assessment in diverse educational settings, such as secondary schools, university courses, and online learning environments, would provide a more comprehensive understanding of its effectiveness and limitations. Specifically, comparative studies could identify how cultural and institutional differences affect the implementation and outcomes of peer assessment.

Finally, experimental studies are needed to test the impact of structured training programs on peer assessors. Research could explore the most effective methods for preparing students to assess their peers accurately and constructively and the influence of such training on the overall quality of feedback provided. By addressing these directions, future investigations can deepen our understanding of peer assessment and optimize its integration into mathematics education.

https://doi.org/10.22521/edupij.2025.14.3 Published online by Universitepark Press

Declarations

Author Contributions. All authors have read and approved the published version of the article.
Conflicts of Interest. The author declared no conflict of interest.
Funding. The author received no financial support for this article.
Data Availability Statement. The data can be provided by the corresponding author upon request.

References

Bedford, S., & Legg, S. (2007). Formative peer and self-feedback as a catalyst for change within science teaching. *Chemistry Education Research and Practice*, *8*(1), 80–92. https://doi.org/10.1039/B6RP90022D

Black, P., & Wiliam, D. (2006). Assessment for learning in the classroom. In J. Gardner (Ed). *Assessment and learning* (pp. 9–25). Sage.

Brindley, C., & Scoffield, S. (1998). Peer assessment in undergraduate programs. *Teaching in Higher Education, 3*, 79–90. https://doi.org/10.1080/1356215980030106

Cheng, W., & Warren, M. (1997). Having second thoughts: Student perceptions before and after a peer assessment exercise. *Studies in Higher Education, 22(2)*, 233–239. https://doi.org/10.1080/03075079712331381064

Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed method approaches.* Sage

Davidson, G. V. (1990). Matching learning styles with teaching styles: is it a useful concept in instruction? *Performance and Instruction, 29*(4), 36–38. https://doi.org/10.1002/pfi.4160290410

de-Armas-González, P., Perdomo-Díaz, J., & Sosa-Martín, D. (2023). Peer assessment processes in a problem-solving activity with future teachers. *Eurasia Journal of Mathematics, Science and Technology Education, 19*(4), em2245. https://doi.org/10.29333/ejmste/13057

Falchikov, N. (1995). Peer feedback marking: developing peer assessment. *Innovations in Education and Training International, 32*(2), 175-187. https://doi.org/10.1080/1355800950320212

Johnson, R. T., & Johnson, D. W. (1985). Student-student interaction: ignored but powerful. *Journal of Teacher Education, 36*(4), 22-26. https://doi.org/10.1177/002248718503600406

Kerr, N. L., & Tindale, R. S. (2004). Group performance and decision making. *Annual Review of Psychology, 55*(1), 623-655. https://doi.org/10.1146/annurev.psych.55.090902.142009

Kirkpatrick, D., & Fuller, R. (1995). The challenge of peer assessment. In L. Summers (Ed.), *A Focus on Learning, Proceedings of the 4th Annual Teaching Learning Forum* (pp. 146–149). Edith Cowan University.

Kollar, I., & Fischer, F. (2010). Peer assessment as collaborative learning: A cognitive perspective. *Learning and Instruction, 20*(4), 344-348. https://doi.org/10.1016/j.learninstruc.2009.08.005

Leikin, R., & Zaslavsky, O. (1999). Connecting research to teaching: cooperative learning in mathematics. *The Mathematics Teacher, 92*(3), 240-246. https://doi.org/10.5951/mt.92.3.0240

Li, H., Xiong, Y., Hunter, C. V., Guo, X., & Tywoniw, R. (2019). Does peer assessment promote student learning? A meta-analysis. *Assessment & Evaluation in Higher Education, 45*(2), 193–211. https://doi.org/10.1080/02602938.2019.1620679

Orsmond, P., Merry, S., & Reiling, K. (2000). The use of student-derived marking criteria in peer and self-assessment. *Assessment & Evaluation in Higher Education, 25*, 23–38. https://doi.org/10.1080/02602930050025006

Radford, L. (2011). Classroom interaction: why is it good, really? *Educational Studies in Mathematics, 76*, 101–115.

Sambell, K., & McDowell, L. (1997). The value of self and peer assessment in the developing lifelong learner. In C. Rust (Ed.), *Improving student learning: Improving students as learners* (pp. 56–66). Oxford Centre for Staff and Learning Development.

Smith, H., Cooper, A., & Lancaster, L. (2002). Improving the quality of undergraduate peer assessment: A case for student and staff development. *Innovations in Education and Teaching International, 39*, 71–81. https://doi.org/10.1080/13558000110102904

Strachan, I. B., & Wilcox, S. (1996). Peer and self-assessment of group work: developing an effective response to increased enrollment in a third-year course in microclimatology. *Journal of Geography in Higher Education, 20*, 343–353. https://doi.org/10.1080/03098269608709377

Strijbos, J. W., & Sluijsmans, D. M. A. (2010). Guest editorial: unraveling peer assessment: Methodological, functional, and conceptual developments. *Learning and Instruction*, *20,* 265–269. https://doi.org/10.1016/j.learninstruc.2009.08.002

Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research*, *68*(3), 249–276. https://doi.org/10.2307/1170598

Topping, K. (2017). Peer assessment: Learning by judging and discussing the work of other learners. *Interdisciplinary Education and Psychology*, *1,* 1–17.

Van Zundert, M., Sluijsmans, D. M. A., & Van Merrienboer, J. J. G. (2010). Effective peer assessment processes: Research findings and future directions. *Learning and Instruction, 20*, 270–279. https://doi.org/10.1016/j.learninstruc.2009.08.004

Venables, A., & Summit, R. (2003). Enhancing scientific essay writing using peer assessment. *Innovations in Education and Teaching International, 40*, 281–290. https://doi.org/1470329032000103816

Webb, N. M. (1991). Task-related verbal interaction and mathematics learning in small groups. *Journal for Research in Mathematics Education, 22*, 366–389. https://doi.org/10.2307/749186

Wen, M. L., & Tsai, C. C. (2006). University students' perceptions of and attitudes toward (online) peer assessment. *Higher Education, 51*, 27–44. https://doi.org/10.1007/s10734-004-6375-8

Xiang, X., Yuan, R., & Yu, B. (2021). Implementing assessment as learning in the L2 writing classroom: a Chinese case. *Assessment & Evaluation in Higher Education*, *47*(5), 727–741. https://doi.org/10.1080/02602938.2021.1965539

Yin, C., Chen, X., & Chang, Y. (2022). Peer assessment as a means to promote language proficiency: Evidence from classroom practices. *Frontiers in Psychology, 13*, 912568. https://doi.org/10.3389/fpsyg.2022.912568

Zariski, A. (1996). Student peer assessment in tertiary education: promise, perils, and practice. In J. Abbott & L. Willcoxson (Eds.), *Teaching and learning within and across disciplines* (pp. 189–200). Proceedings of the 5th Annual Teaching Learning Forum, Murdoch University, February 1996. Perth: Murdoch University. Retrieved April 1, 2008, from http://cea.curtin.edu.au/tlf/tlf1996/zariski.htm

Zevenbergen, R. (2001). Peer assessment of student constructed posters: assessment alternatives in preservice mathematics education. *Journal of Mathematics Teacher Education. 4*, 95–113. https://doi.org/10.1023/A:1011401532410

About the Contributor

Yaniv Biton, PhD, Shaanan Academic College, Center for Educational Technology, Tel Aviv, ISRAEL
Email: YanivB@cet.ac.il