MODESTUM

EURASIA Journal of Mathematics, Science and Technology Education, 2025, 21(7), em2667
ISSN:1305-8223 (online)

OPEN ACCESS

Research Paper

https://doi.org/10.29333/ejmste/16605

Student and teacher learning as a result of developing peer assessment criteria for mathematical tasks

Yaniv Biton [1,2*]

[1] Shannan Academic College of Education, Haifa, ISRAEL
[2] Center for Educational Technology, Tel Aviv, ISRAEL

Received 01 May 2025 ▪ Accepted 22 June 2025

Abstract

The idea of involving students in the assessment process has recently taken a prominent place in education. Research suggests that peer assessment in which the students define assessment criteria and assign value enhances learning. The study presented herein is part of an ongoing research project being conducted at a pre-academic course at a technological university in which students are tasked with developing and applying criteria for grading mathematical exercises. The aim of the study is to explore the potential of incorporating such peer assessment into learning mathematics. We also investigate the teacher's reflections regarding such a peer assessment activity. The findings suggest that when students are actively involved in developing criteria and use them to grade other students' work, they see 'behind the scenes' of the mathematical problem, which improves their understanding. The findings also suggest that the teacher's reflection on the process helps better refocus on lesson plans.

Keywords: peer-assessment, learning, developing criteria, pre-academic study, higher education

INTRODUCTION AND THEORETICAL BACKGROUND

Assessment to Enhance Learning

Recent years have witnessed a shift away from 'assessment *of* learning' (i.e., for the most part, summative assessment) and a trend towards 'assessment *for* learning', meaning using assessment tasks to promote learning (Cartney, 2010; Keppell & Carless, 2006). Involving students in assessment processes, specifically, peer assessment, in educational settings has been gaining increasing prominence.

Involving students in assessing their peers can provide them with skills to judge what constitutes high quality work, thus increasing their cognitive and mega-cognitive understanding (Topping, 1998). Peer assessment affords an opportunity to see examples of how others solve problems and approaches that differ from their own and provides them with the ability to learn by internalizing given criteria and standards (Gielen et al., 2010). It can be applied to evaluate specific task outcomes, individual contributions, or collaborative group efforts (de-Armas-González et al., 2023; Kollar & Fischer, 2010; Topping et al., 2020).

Formative vs. Summative Assessment

Peer assessment may serve formative and/or summative purposes. The advantage of formative assessment (not necessarily peer assessment) is to allow early identification of errors in thinking and misconceptions, thus facilitating intervention that can lead to deeper understanding and knowledge retention and increase student learning. This contrasts with summative assessment, which only confirms or negates learning (Keppell et al., 2006).

Formative evaluation can focus on cognitive, social, affective, and/or meta-cognitive aspects. It generally includes a multi-method approach to create a complete profile of the student instead of a single score. Strijbos and Sluijsmans (2010) noted that because peer assessment involves reflection, discussion, and collaboration, it also serves as a formative evaluation for the assessors and may be seen as an integral part of the learning process that takes place during a learning program and not only at its end.

✉ yanivb@cet.ac.il **(*Correspondence)**

Contribution to the literature

- This study expands current research on peer assessment by exploring its structural components and emphasizing the significance of students developing assessment criteria on their own.
- The study sheds light on the added pedagogical value of engaging students in constructing and applying evaluation benchmarks, highlighting both the learning opportunities and the inherent challenges.
- The findings also emphasize the professional learning of teachers through their reflections on students' assessment processes, offering insight into how such experience can inform future instruction and planning.

What is 'Peer Assessment'?

Topping (1998) defines 'peer assessment' as a process in which students judge their peers quantitatively (supplying them with scores and grades) and/or qualitatively (providing them with written or oral feedback). It can be a valuable addition to enhance students' learning for both the assessors and assesses.

Peer assessment *for* learning has three main elements:

(1) assessment tasks that focus on the concepts under study (Black & William, 2006; Boud, 1990),

(2) student participation in designing and performing the assessment process, which requires judgement and reflection (National Council of Teachers of Mathematics [NCTM], 2009; Orsmond et al., 1996), and

(3) the transmission of feedback between peers and teacher to increase learning, provide formal accountability, allow accreditation of the knowledge gained by students, and make each aware of the others' goals (Orsmond et al., 2000).

Developing Criteria

An important stage of the assessment process is developing the criteria. Orsmond et al. (2000) defines a criterion as '... a distinguished property or characteristic of anything, by which its quality can be judged or estimated, or by which a decision or a classification may be made' (p. 30). Strachan and Wilcox (1996) consider the stage of developing criteria a critical one.

Typically, assessment criteria are predefined to align with the formal requirements specified for the evaluated task. Supplying students with marking criteria benefits the students as they can use this information to assess their own work. Enabling students to monitor their own learning rather than relying on their teachers for feedback can help them become realistic judges of their own performance (Sambell et al., 2006).

However, going through the process of peer assessment and–even more so–having to define criteria on their own, can substantially improve the learning experience. As a case in point, in a study by Williams (1992) regarding peer assessment, students used teacher-prescribed rubrics to assess the work of their peers. The students enjoyed the activity because they felt it benefitted their learning but noted that they would have liked to have a greater say in designing the assessment process and to have received guidelines and training for their role as assessors. Orsmond et al. (1996) also explored the issue of marking criteria. In earlier studies, students were supplied with assessment criteria (Orsmond et al., 1996, 1997a, 1997b). However, a later study suggested that allowing students to form their own criteria provided greater benefits (Orsmond et al., 2000).

When students take an active role in designing the assessment process through negotiations regarding the criteria, the procedure, or its interpretation, they become better acquainted with the concepts under study. Developing criteria can be done by the students on their own or with help from their teacher. In fact, active cooperation between teacher and students, whereby teachers give detailed instructions to students on setting criteria, and the student and instructor clarify objectives and standards together are important parts of the peer assessment process (Somervell, 1993; Strachan & Wilcox, 1996). When educators and students jointly construct criteria for marking, the students' understanding of the marking criteria are enhanced (Brown & Knight, 1995). Furthermore, using negotiations on the basis of shared meaning has been found to increase student-student and student-teacher interactions and can be used to enhance learners' understanding of other students' ideas during the learning experience (Falchikov, 1995, 2001; Moschkovich, 1996, 1998).

In the research described herein, groups of students (two or three students in a group) developed criteria independently without any intervention by the teacher. Later, the teacher followed the proceedings by viewing the audio-video documentation and used this information to draw conclusions about her teaching in the classroom and deciding if the lessons needed augmentation. This is also an important aspect of peer assessment (how it affects the teacher) and is also discussed herein.

Additional Aspects to Enhance Peer Assessment

According to Topping et al. (2020), peer assessment generally results in a significant improvement in 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 conduct the assessment, the effect size was markedly larger, suggesting that structured training enhances both the accuracy and educational benefit of peer assessment. Additionally, digitally conducted peer assessments was 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 that incorporate structured training and that leverage digital tools to maximize learning outcomes (Topping et al., 2020).

Benefits of Peer Assessment

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).

Recent studies have further explored the mechanisms of peer assessment, particularly within educational contexts focused on developing critical problem-solving skills. For instance, de-Armas-González et al. (2023) investigated pre-service secondary school mathematics teachers who engaged in assessing their peers' solutions to open-ended mathematical problems. Their study revealed that although the process of assessment is often influenced by the assessors' own problem-solving approaches, something that can challenge the objectivity of evaluations, peer assessment activities helped 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.

Biton (2025a, 2025b) provides a comprehensive exploration of how peer assessment can serve as a catalyst for mathematical learning, particularly in preparatory academic settings. His studies focus on the dialogic and collaborative processes that emerge when students evaluate each other's solutions to complex mathematical problems. Uniquely, the studies emphasize the formative potential of student-led development of assessment criteria and examine the learning opportunities embedded in peer interactions, including the ways in which students identify errors, justify their reasoning, and refine their own understanding. The findings demonstrate how assessing peers' work not only enhances conceptual understanding but also fosters mathematical discourse and metacognitive awareness. The studies offer empirical support for the value of peer assessment in promoting student agency, reflection, and depth of mathematical thinking.

Encouraging students to evaluate each other's work, fosters active participation and reflective learning. Therefore, introducing a peer-assessment process into the learning paradigm can add an additional instrument for learning as it ultimately aims to make students plan their own learning, discover and define their weaknesses and strengths, and improve their transferable skills (Boud, 1990; McLaughlin & Simpson, 2004).

Study Goals

In our study, we focused on the stage in which a group of students sit together and use open discourse to discuss the criteria needed to assess a particular mathematical problem. Communication in the group was oriented towards understanding and respecting the perspectives of others, based in part on Taylor et al. (1997):

> Open discourse gives rise to opportunities for students to:
>
> (1) negotiate with the teacher about the nature of their learning activities;
>
> (2) participate in the determination of assessment criteria and undertake self-assessment and peer-assessment;
>
> (3) engage in collaborative and open-ended enquiry with fellow students, and
>
> (4) participate in reconstructing the social norms of the classroom (Taylor et al., 1997, p. 295).

As will be shown, the process of formulating criteria included multiple interactions among the members of the groups. This is important because, as noted by Saxe et al. (2009), it facilitates the travel of ideas by requiring each student to both present and listen, thus enhancing their engagement in the process. This can also be considered an aspect of 'collaborative learning' or 'peer learning', based on the definition '... the use of teaching and learning strategies in which students learn with and from each other without the immediate intervention of a teacher' (Boud et al., 1999, pp. 413-414). In this sense, therefore, peer assessment may be considered a subcategory of peer learning.

To summarize, this study aimed to explore potential learning opportunities when students are tasked with developing criteria for peer assessment in mathematics. While the main focus is on student learning, the teacher was actively involved in all stages of the process, including selecting tasks, facilitating decision-making, and analyzing the outcomes. We examined both student and teacher learning.

Research Questions

1. What learning opportunities emerge for students during the process of developing peer-assessment criteria in mathematics, and what types of learning do they demonstrate as a result?
2. What learning opportunities are created for the teacher as a result of observing and engaging with the students' process of formulating peer-assessment criteria?

MATERIAL AND METHODS

Research Locale and Participants

The study took place at the Center for Pre-University Education of the Technion–Israel Institute of Technology. The study group comprised approximately 30 students (two classes) who were participating in an intensive, single-semester preparatory course in mathematics to bring them up to level-five matriculation (highest possible level) (similar preparatory courses are offered in many universities worldwide to close any gaps between students' achievements in high school and prerequisites for academic study). One teacher (f.) taught both classes. The study took place in Hebrew.

Typically, students' progress was assessed systematically using traditional methods (i.e., exams), but during this course, students were given an assignment to work in groups to assess their peers' work. Six peer assessment activities (three from each class) were recorded using a video camera and audio recordings.

Peer Assessment Activities

Each of the six peer assessment activities was 90 minutes long and comprised of three stages.

Stage 1 (25 minutes)

Students individually solved a mathematical problem that the teacher had chosen from the course textbook. A sample problem:

> Given: A right quadrilateral pyramid with a square base. The height of the lateral face of the pyramid is $\sqrt{6}$ cm.
>
> A. Find the length of the side of the square for which the volume of the pyramid attains its maximal value.
>
> B. Find the maximal value of the volume.

Stage 2 (15 minutes)

We randomly divided the students into groups of two or three. Each group worked together to formulate criteria to evaluate the solution(s) to the problem and to weigh the criteria according to their relative importance. Each group received a four-page booklet in which to write their evaluation criteria on the first page and used the remaining pages to comment and score the other students' solutions. **Figure 1** presents an example of one group's peer assessment page, showing their criteria for assessing the solution to Item A in the above problem (first column), how they weighed the criteria (second column), the actual scores they gave to one assessee for each of the criteria (third column), and the assessors' reasoning for the scores (fourth column).

Criterion	Maximal score	The actual score	Justification
Theorems and Proofs	30	0	No justification related to solid giometry
Construction of the function	30	30	correct function
Calculation	10	10	
Derivative and Substitution	30	20	The derivative is wrong.

Figure 1. Example of assessment page (Source: Field study)

Stage 3 (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). Inter-group conversation was allowed and encouraged at all stages of the activity.

We advised the students that this was merely an exercise, and their assessment would not affect their peers' final marks. This was to prevent them from feeling inhibited about giving a low mark or a critical 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.

Data Collection and Analysis

Data collection combined multiple modalities to allow rich, triangulated analysis.

Audio-visual recordings

Three video cameras and four audio recording devices were strategically placed to capture both verbal interactions and written actions during the peer-assessment stages. The video documentation was critical for analyzing the dynamic discussions, gestures, and referencing to written materials. The recordings were transcribed in full, and sections relevant to key decision-making moments in groups were analyzed using discourse analysis techniques, enabling insight into students' reasoning, negotiation patterns, and emerging understanding. The video data were analyzed through an iterative process of open and axial coding, using a grounded theory approach to identify patterns of interaction and emerging themes. Special attention was given to sequences of peer negotiation, justification practices, and spontaneous use of mathematical language. Segments were indexed according to cognitive acts (e.g., explaining, questioning, disagreeing) and were cross-checked by two independent researchers to ensure coding reliability and interpretative validity.

Written documentation

All students' written products–initial individual solutions, criteria formulation pages, and peer feedback sheets–were collected for systematic qualitative coding and thematic analysis.

Cross-referencing

Cross reference of the transcripts with the written artifacts allowed the researchers to track how the students' initial attempt at the solution influenced their suggestions for criteria and scoring. This multi-layered data analysis supported the identification of both cognitive and social mechanisms underlying peer assessment and illuminated the opportunities for individual and collaborative learning inherent in this activity.

RESULTS

To allow for an in-depth analysis and to illustrate the rich cognitive and social dynamics embedded in the activity, this paper focuses on a single case study of a complete peer-assessment process–from the formulation of evaluation criteria to the final scoring–in detail. It illustrates how one group of students (Benny, Lior, and Tamir) designed their peer assessment for the problem of extreme value in the field of spatial trigonometry given above. This includes excerpts of a prolonged discussion that involved formulating criteria, naming them, and allotting scores (stage 2). An important aspect we examined is how their own solutions in stage 1 affected their ideas for stage 2. The excerpts also include examples of how disputes were resolved.

Negotiations on the Way to Formulating and Grading Criteria

After stage 1 was completed, stage 2 was when the students first sat together to jointly formulate criteria for evaluating the problem and to decide on maximum points for each. Viewing the documentation of the interaction in the group, we found that during the process of formulating the criteria, the students were often divided in their opinions about which criteria were appropriate for evaluation, how to label each criterion, and what percentage of the score to allocate to it (in relation to its importance), and so forth. We observed that accompanying the disagreements, the students endeavored to give reasons to justify their stance and convince each other.

Excerpt 1

The process of criteria formulation began with Benny's suggestion for a criterion that he suggested calling 'the method'.

> Benny: 'The method'.
>
> Lior & Benny: The manner in which the work was done.
>
> Lior: If he [the problem solver] understands the problem.
>
> Benny: If he understood the point of the problem.
>
> Tamir: I think 'the method' is too general. The method should be divided up into stages. For example, did he define the volume correctly? And then find the correct derivative.
>
> Lior: We need to make sure he explained the stages he used and clearly understood the mathematical procedure. This is the method of doing this work.
>
> Benny: First, if he at all understood how to determine the parameters of the problem.
>
> Lior: What was given in the problem.
>
> Benny: So, we won`t write 'the method'. It's too general. First thing should be if they determined the correct parameters of the problem and defining it. Which parameter?

As can be seen, Benny suggested a criterion that he labelled 'the method'. Lior at first agrees, but Tamir's intervention helps them both understand that this is too general and suggests dividing it into more concrete ones, such as function of volume, the derivative, and establishing parameters. We can see that the other two accept Tamir's suggestion: they reiterate Tamir's words and divide 'the method' into sub-criteria.

To better understand the considerations behind the suggestions, we examined their solutions to the problem in stage 1 (before formulating criteria). **Figure 2** and **Figure 3** show Benny's and Tamir's solutions, respectively. We were surprised to find that Benny, who had solved the problem in detail and given verbal

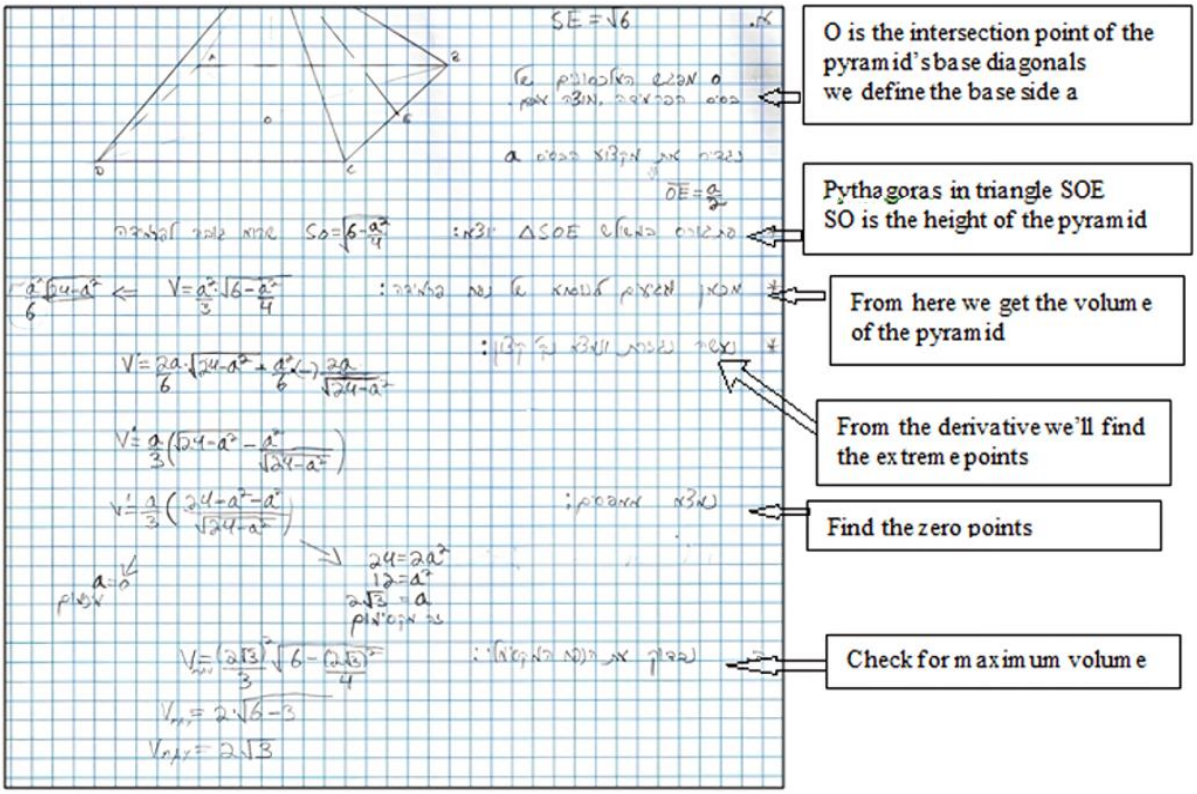

Figure 2. Benny's solution (Source: Field study)

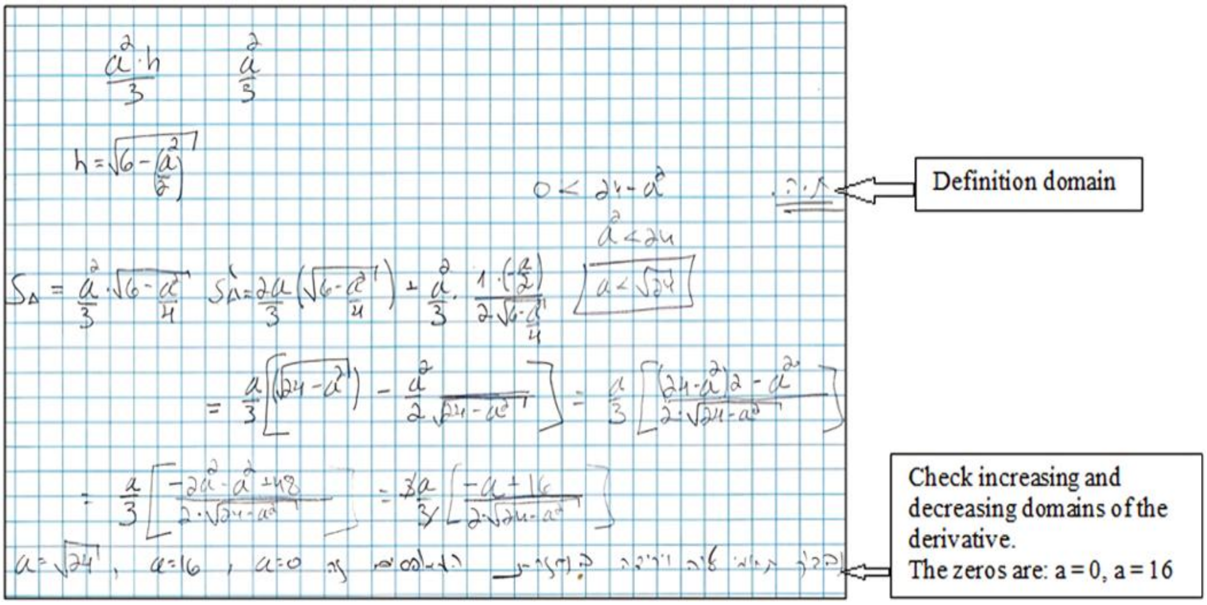

Figure 3. Tamir's solution (Source: Field study)

explanations, was the one to suggest such a general criterion as 'the method'. Tamir's intervention, claiming that this criterion was too general, caused Benny to review how he solved the problem and then to suggest 'First, if he at all understood how to determine the parameters of the problem'. If we review the first and last lines of the discussion, we see that he did understand that his suggestion was too general and that the criterion should be divided more concretely.

Excerpt 2

We next consider the discussion concerning the criterion 'establishing parameters'. At this point, they have reached the stage at which they want to decide the weight of the criterion. As can be seen from the following excerpt, an argument is tossed back and forth between Benny and Lior concerning the importance of the definition of the parameter vis-à-vis its contribution to the furtherance of the solution.

Lior: Let's discuss 'establishing parameters of the problem'.

Benny: If he defined the correct variables… how many points should that be?

Tamir: Five.

Lior: I agree. Five points.

Benny: Will you have other criteria to add up to 95 more points?

Tamir: There are more [criteria that will be allotted points].

Benny: I think it should be more. More. Eight is better.

Lior: Look. He wrote that x is the length of the base [argues that this criterion does not deserve more than 5 points].

Benny: So, he knows that this is the problem. He understood that this is the variable needed here. He correctly understood the essence of the problem, that is, that one needs to find that variable. That's the problem.

Lior: He understood that this is the unknown. I think 5 points is enough.

Tamir: You know what, don't write that now. Let's wait.

All talking together: We'll decide the points at the end.

It is interesting to observe the arguments given in the group during the process and to relate the arguments to the personal experience of each when they solved the problem in stage 1. Examining Benny and Tamir's solutions (**Figure 2** and **Figure 3**), leads us to surmise that Tamir's wish to allocate only five points to this criterion emanates from the fact that he did not at all mention the establishment of the parameter in his own work and therefore feels that it is not worthy of too many points. This contrasts with Benny, who did consider this in his solution and that is probably why he thought it should be allocated more points.

It is also interesting to observe that the students did not relinquish their right to determine and influence the decisions made. They clarified their stances about what they saw as more or less important to the process of finding a solution and attempted to convince the rest of the group of their view. For the most part, the group arrived at a mutual agreement, and in the case above, the disagreement ends with the group postponing any decision regarding the points for each criterion until they have established the full list. As we can see, the situation of reaching decisions in a group confronts the assessors with further challenges beyond that of finding a correct mathematical solution.

Excerpt 3

As the process of consolidating the criteria continued, Tamir suggested that one criterion should be 'function's domain'. In the following excerpt we can see that Benny and Lior disagree with Tamir's suggestion and refuse to write it down as a criterion.

Tamir: Hm … we didn`t add 'function's domain'.

Benny: No need.

Tamir: Sure there is!. x can't be greater than 6.

Lior: That is a solution of the function or the derivative. Forget function's domain.

Possibly the reason for the refusal to add 'function's domain' as a criterion arises from the fact that Benny and Lior both did not consider it to be a stage by itself in their solutions, in contrast to Tamir who did consider it as such (see **Figure 3**). The argument continued with reasons for and against.

Tamir: If you don't add 'function's domain', then your function, your derivative, continues [is defined] for all x.

Benny: Here you can [have] 'x's which are greater than or equal to 6.

Tamir: No … the square root of x is greater than the root of 24 [explains what he had written in his solution].

Benny [addressing researcher]: Do we need 'function's domain'? [note: researcher did not respond].

Benny [to group]: This isn't [an investigation of] a function. A function's domain is needed if you find points and they aren't in the function's domain, then you take off points for this.

Lior: Benny is right. You don't need 'function's domain'. You need to check afterwards if your solution is within the function's domain.

Tamir: Okay, but why …

Lior: Because you weren't asked to investigate it.

Tamir: This is not something that needs grading. It's something that the solver needs to understand to solve the problem. But it isn't something that he is asked to define as part of the problem.

Table 1. Benny, Lior, and Tamir's list of criteria for assessing the problem

Criterion	Points
Correct drawing of the problem	10
Establishing correct parameters of the problem	5
Expressing the variables through the parameter of the problem	10
Finding a function of volume by using parameters and correct substitution	5
Finding a correct derivative for the problem	30
Finding the maxima and minima points	25
Choosing a correct point	15

Lior: Okay. So, find maxima and minima points and then check if it's within the domain.

Benny: [I don't think it's important enough] that you have to take off points from those who didn't write it.

Similar to the previous examples, here it also seems that the considerations for choosing a criterion and points for it are influenced by the personal experience each assessor had with the problem in stage 1, and that is why both Lior and Benny insist that there is no need to define the domain. But note that this insistence does not necessarily arise from them being convinced this domain is unnecessary. Actually, we have evidence of their feelings from their own personal experience. At the end of the discussion, we can see that the disagreement is resolved through a kind of 'power game' in which the majority decides.

Final List of Criteria

Table 1 shows the final list of criteria formulated by the students for the evaluation of part A of the problem. Note that if the second, third, and fourth criteria are combined, this in fact constitutes the construction of a volume function and together receive 20 points. This distribution of points differs from that of the Ministry of Education's. As the teacher later explained: 'The Ministry of Education states that in questions of minimum and maximum, 40% is for constructing the function, and 60% is for the solution… meaning that they [the students] assigned less importance to constructing the function than the Ministry of Education' (later, we will discuss what the teacher has learned from observing the students' criteria and their discussion regarding the formulation of this criteria).

DISCUSSION

Although we have presented only one case study for our findings, our observations of all the groups suggest that peer assessment activities help students better understand the mathematical concepts being taught as a result of having to consciously formulate the stages and requirements for solving mathematical problems, discuss them with their peers, and decide on which aspects of solving the problem have more or less importance. The process is also valuable for the teacher, who, by observing the students' performance, can deduce whether what was taught in class was sufficient for the students' understanding of the concepts.

Designing Peer Assessment Actions

We noted that the evaluation criteria that the students proposed were related to how they themselves had solved the problem and what they believed was important in the solution process. After completing stage 1 (individual solutions), the students were able to suggest one or more criteria that they felt were important in solving the problem, which were mostly a result of them trying to re-create the stages they took in solving the problem.

This was followed by a process of group decision-making where the students decided whether to accept the suggested criterion. While observing the dynamics of the groups, we noted that the way in which a student designed their peer assessment actions affected the way they engaged in the group 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, alongside their role as learners, required them to provide well-structured explanations and ask questions; that is, the need to learn about the correct solution to the problem and to assess it appropriately. These situations evoked discussion and active listening throughout the phases of the assignment. Choosing the students for each group at random often introduces heterogeneity, and this seems to have 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 an 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).

After consolidating their ideas regarding criteria, the next action the students took was to assess the solutions of the students in the other groups. We noted that this sometimes led them to alter the criteria or the weight they had assigned, a finding that aligns with what is known from the literature about the ability of assessors to exercise discretion concerning the relationship among the product, the goals of the task, and criteria for success (Saxe et al., 2009).

Challenges in Integrating Peer Assessment Activities

Group heterogeneity can be an effective tool for enhancing collaborative learning (Davidson, 1990; Johnson & Johnson, 1985). Similar to other studies on collaborative learning, we observed that when groups included students who succeeded in solving the assignment alongside those who could not, heterogeneity paid off. However, heterogeneity is liable to introduce challenging issues, the top four being the influence of group pressure, advanced students, less motivated students, and fear of causing distress to the classmates. We aimed to overcome these challenges as described below.

First challenge: Group pressure

As part of a collaborative group, the student needs to feel throughout that they are a member of a team. However, they may feel pressured by the group and as a result feel inhibited in expressing themselves by suggesting an idea that deviates from those given by others.

We attempted to mitigate this by having the students first solve the assignment independently (stage 1). This, we felt, would contribute to the discourse as each would have first consolidated their own ‘position’ concerning how to solve the problem, which would, hopefully, give them more confidence to contribute to the discourse. Also, if they encountered difficulty when solving the problem, the group interaction would allow them to express and deal with those difficulties, to observe those encountered by others, and to engage in brainstorming. This proved valuable later when it came to setting the criteria for assessing the solutions (stage 2) as well as when discussing the correctness of the solutions they assessed (stage 3).

Second challenge: Advanced students

The need for conformity in a group can be a disadvantage for advanced students because it reduces their average to that of the group and can limit their thinking skills and creativity. However, we observed that the process actually enhanced learning for successful students (including those who only partly succeeded), because during the group discourse, they very quickly took on the role of ‘teacher’ to answer their peers’ questions. This situation is well known and supported in the literature: the act of explaining serves a distinct advantage for the ‘teacher’ as it reinforces their knowledge, can shed light on aspects which they themselves might have been previously unaware of, give them better understanding of the concepts, and improve their ability to articulate those concepts that were unclear to their peers (Webb, 1991).

Third challenge: Less motivated students

It is a common phenomenon that in a collaborative setting, weak, lazy, or uninterested students will tend to relinquish their responsibility for their own learning and attempt to benefit from their collaboration with more diligent ones.

During our pilot study, various options for conducting the peer assessment tasks were tried. One of them was a situation where the students were divided into pairs: one solved the problem and the other assessed it. In this case, we noticed that students who failed to solve the problem on their own (i.e., less motivated students) refrained from taking part in the one-on-one peer assessment, claiming that they could not assess a solution they had failed to find themselves. This phenomenon of social avoidance during collaborative learning is well known in the literature (Kerr & Tindale, 2004). For this reason, we altered the paradigm to that of a group (instead of a pair) and in this present study did not observe this phenomenon. In fact, we noticed that students who initially had difficulty solving the assignment in stage 1 still took an active part in the subsequent stages. That is, they asked questions, made attempts to probe deeper and to understand the phases of the solution offered, and took a stand on the validity of various criteria, even voicing their opinions concerning the grade they had to give.

Fourth challenge: Fear of causing distress to their classmates

When peer assessment is done publicly, the assessors may fear that they will cause distress to their peers or have difficulty being objective due to social concerns. These are well-known phenomena recognized in the literature. Indeed, the participants in our study did initially express concerns about their assessments influencing their peers’ final grades, but we assured them that the grades they assigned would not influence the formal ones. This led them to be more relaxed and initiate assessor-assessee meetings spontaneously.

Analysis of Learning Opportunities Resulting from the Process of Formulating and Grading Criteria

Requiring each student to first deal with the problem individually allowed them to consolidate their own ideas and methods for solving their problem and then share with the group the ideas, feelings, and difficulties which had arisen during the individual experience. The next step, which was to establish criteria, gave the students the authority to decide how to evaluate the problem and fostered attentiveness and discussion during the execution of the task. It made them aware of the other group members’ understanding. This, and the fact that the groups were small (only 2-3 members) and heterogenous in terms of abilities and personalities generated a robust exchange of ideas (cf. Orsmond et al., 2000). The heterogeneity existing among the students in each group was extended by the difference in each one's personal experience with the problem. Bringing together learners from diverse backgrounds and with different abilities enables each to express their experience and contribute from it, all of which meet with the principles of co-learning. From the dialogues presented above, it is apparent that the choice of criteria was mostly based on how the students had arrived at their solution. The experience of solving the problem created a plethora of opinions about the components of the problem and their relative importance, which influenced the process of decision making in the group.

In the course of negotiating what evaluation criteria to use and the weightings, the students frequently disagreed with each other. The discussions were not without conflict and disagreements led to fervent conversations where each attempted to vindicate their position. All the students defended their positions and did not abandon their right to influence the decision-making. Although the arguments were sometimes intense, most resulted in reaching a consensus based on mathematical considerations and not on social dynamics (cf. Saxe et al., 2009).

Not every suggestion ended in agreement among all the members of the group regarding the need for a certain criterion. For example, in excerpt 1, Benny suggested a criterion called 'the method' and in excerpt 3, Tamir suggested a criterion called 'function's domain'. Both suggestions led to disagreement. If we compare the nature of the two discussions, the former ended without an argument, whereas the latter created a situation of uncertainty and imbalance (cf. Harel & Koichu, 2010) that impeded continuation of the process. resulted in a rather heated dispute about whether to include that criterion. In the first example (excerpt 1), Benny`s suggestion came after an internal process in which he searched for, and conjectured about, a criterion which would be suitable for assessing the problem. Benny clarifies his intent later, resulting in a chain of reactions entailing further mental acts, such as exposition and a suggestion to search for another criterion. We can see that the group entered a mode of imbalance, although short-lived, in which they understood that Benny`s suggestion was too general, leading to a need for another, more specific criterion under another name.

In the continuation of excerpt 1, Tamir's suggestion led Benny to reflect on the way he had earlier solved the problem and to suggest a new criterion. In effect, this reflection signals the group's emergence from the state of imbalance: the understanding achieved by the assessors as a result of the discussion led to the need for a new criterion which is, in fact, one of the first stages in solving the problem. These 'mental acts' exercised during the discussions among the group members suggest that students may reach desired understanding by the application of mental acts such as deduction, generalization, construction, classification, searching, explication, application, and prediction, resulting in the learner constructing mathematical products.

In comparison to excerpt 1, which ended without argument, in excerpt 3, Tamir`s suggestion for the 'function's domain' criterion drew the students into a state of uncertainty concerning the need. In the first phase, an attempt was made to reject that suggestion because two of the assessors did not take it into account in their own solutions. However, in contrast to excerpt 1 (where Benny did not insist on 'the method' criterion offered by him), here Tamir does not give up and is convinced he is right. It appears that all attempts to reject his suggestion cause him to search for additional explanations to prove the need for a 'function's domain'. Further evidence for the state of uncertainty can be seen when Benny referred the question regarding the need for 'function's domain' to the researcher. An explanation for this can be found in Piaget's (1977) theory, that is, when a person feels tension, it creates motivation to act in order to return to a cognitive balance.

Also interesting is Lior's attempt to strongly reject the suggestion for 'function's domain', starting with 'leave the function's domain', continuing with words urging the move to another criterion, and ending with Lior's consent following Benny`s explanation why it was possible to relinquish 'function's domain'.

Tamir is unable to convince the others that this criterion should be considered. However, after repeated viewings of the documentation, we can discern that Benny's and Lior's voices broadcast insecurity, and they are careful with the explanations they offer against adding that criterion. This is in contrast to Tamir, who tries to justify his choice of that criterion.

Grading

The attempt of all three to formulate and grade criteria simultaneously failed (excerpt 2), and a decision was reached only at the end of the task. Explaining, questioning, and reflecting on how they had solved the problem themselves led to a discussion that helped them exit the state of imbalance and move on to the stage at which the students graded the criteria (cf. Harel & Koichu, 2010). This may be considered an additional reflection stage during which they reviewed the difficulties they encountered during their initial search for a solution and succeeded in explaining to their colleagues the reasons for their grading. It is interesting to see the imbalance created among the assessors' opinions. On the one hand, it is their wish to grade the criteria according to importance but on the other is their wish to keep the grading according to the stages they themselves had used in solving the problem successfully.

Opportunities for Teachers

Even though the goal of the present research was to follow the students' learning process emanating from peer assessment, it was difficult to ignore the process of learning undergone by the teacher. As a case in point, the teacher pointed out that sometimes groups (not only that of the case study) made decisions about criteria and grading that were not precise. This suggested to her that the students seemed to have an inaccurate impression of what specific criteria should be. These situations, known in literature as learning inhibitors (Balacheff, 1991), offer learning opportunities for the teacher, who can form a better appreciation of what the students understand (cf. Norton et al., 2011). For example, the criterion 'function's domain' was not included in the final list of criteria, and there were situations where the students allotted greater or less weight to the criteria than what the standard curriculum dictates, based on what they deemed more or less important. However, these and similar situations allowed us to discern opportunities that are important for teachers. As we noted, the teacher was attentive to the students' discussions and questions regarding the names of criteria and their grading. These are observations that the teacher can later consider when preparing lessons.

The teacher also observed the video documentation and discussed the findings and conclusions emerging from each task with the researcher. The following is an excerpt from a conversation with the teacher regarding the list of criteria formulated in the group (see **Table 1**).

> Teacher: They [the assessor students] allotted less importance to constructing the function than the Ministry of Education does. I guess that is because their criteria are based on what is easy or difficult for them and they probably consider this part to be easy. I will certainly talk with them about this … because later they constructed an incorrect function, explored it, and then said on the basis of the function they had constructed that everything is correct.

This excerpt suggests that the teacher's exposure to the students' criteria and their grading enabled her to discern what the students consider more and less important. In addition, on the basis of this list, she learned that the students probably do not understand the overall importance of the problem. It appears that the students did not 'stretch' their thinking but confined themselves to what they felt comfortable or secure with, that is the knowledge they already had (cf. Orsmond et al., 2000). The teacher continued:

> The questions that the students asked during the discussion were meaningful. The students must understand that the criterion and its point allotment should not be based on level of difficulty but on level of importance. The grading they allotted shows me that they don't understand the importance of the exercise … this really requires attention.

Later, after reviewing the documentation more closely, the researcher returned to the teacher with different situations, among them the 'function's domain' situation. Here is her reaction following their joint review of the video documentation.

> This is invaluable. This contributes greatly to my insight. The fact that 'function's domain' came up and the ensuing argument … is an angle that does not arise in frontal teaching. Look, I can tell them again and again about 'function's domain' [in my lessons], but they get so much information in the Pre-Academic Center to absorb that it's impossible to expect each and every one of them to grasp how important identifying the domain of a function is when solving an optimization problem … The fact that they are talking about it is really, really good. Considering the domain as a criterion shows that they have a deep understanding of the structure of the problem.

The teacher's comments suggest that the peer assessment activity has the potential to strengthen and complement aspects that cannot be achieved with frontal teaching and reinforces the literature about the value of peer feedback (in this case, within the group), especially that feedback given by peers was more significant for students than feedback given by an expert (i.e., the teacher) (Cho & MacArthur, 2010).

A more concrete example of how the 'function's domain' criterion enhanced the teacher's learning can be seen in the following:

> I notice that they haven't paid any attention to the criterion of 'limited domain' … This suggests to me that I probably didn`t stress it sufficiently in class. If nobody talked about it, it means that I didn't pay enough attention to it. I will have to review it with the students to draw their attention to it.

Aligning Student-Teacher Expectations

Another aspect we investigated was how such an exercise would help coordinate the expectations between teacher and students. The students, for their part, saw the process of choosing and grading the criteria, in general, alongside the evaluation document in particular, as two important components of such. As one student remarked:

> I think it will benefit the teachers if they see the student's action. Because what I have done here, I have given the kind of output I expect to get, or

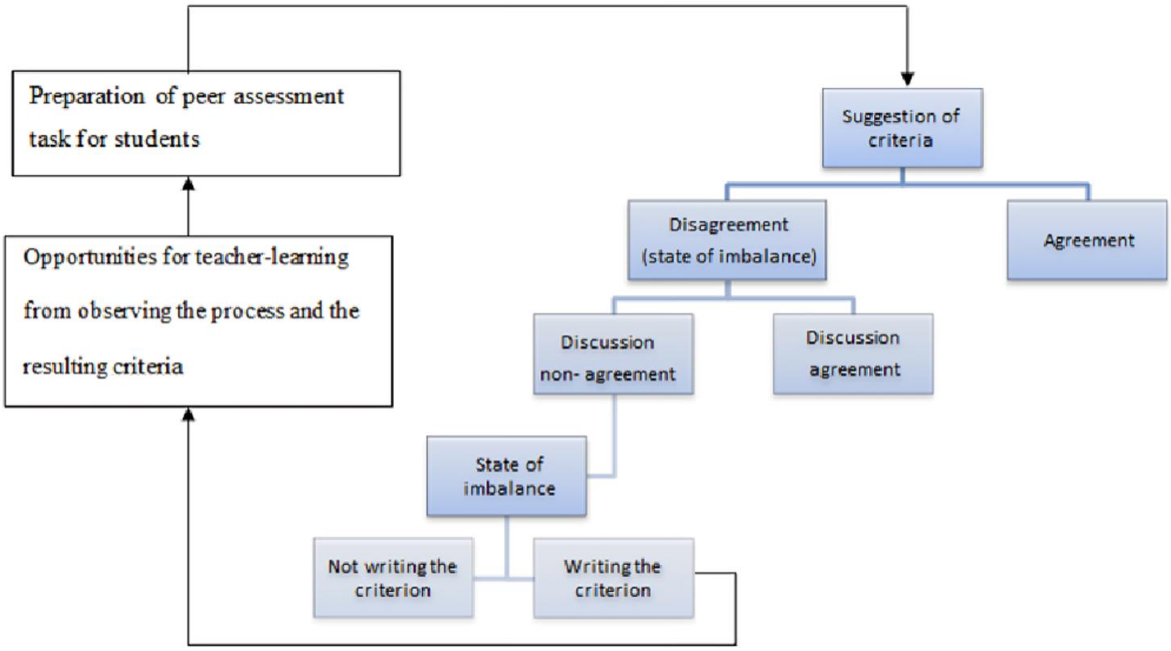

Figure 4. The process of formulating a criterion for assessing a problem and the opportunities for learning arising from this process (Source: Author's own elaboration)

> what seems appropriate to me. I think it will help the teacher if she sees what I see. This seems to me an advantage, to kind of coordinate between the teacher and student, each [saying] what they expect from the other and from themselves … This instrument is wonderful in that it allows the student to express themself and enables the institution to look at it and reflect on it. This is a window [of opportunity] for the student to talk.

The case study we presented confirms how the students related to the stage in which they were required to formulate criteria. The interviews and questionnaire illustrated that the students comprehended that the stage of formulating the criteria was significant for their learning, as can be seen from the following two statements about formulating the criteria taken from the interviews.

> Student A: To sit and think together on what [basis] to define the criteria … to be able to understand what stands behind every problem.

> Student B: Suddenly, you look at a question and think, 'How will the teacher divide the points?' And then, according to this, you also solve [the problem].

In other words, in this evaluation approach, as the learner considers how to design the assessment, they are exposed to techniques and processes carried out by the teacher (cf. Zariski, 1996) and this process enables them to see the problem from the teacher's point of view. In the future, they will solve problems by thinking of their components. This also contributes to improving the teacher's point of view as it affords them the opportunity to discover which components of the problem are not obvious to the students and thus prepare them for situations where they have to solve similar problems under pressure, as in a test.

> When students gain experience in understanding the various aspects of problem solving–that is, which are more or less important–they are better able to deal with solving problems when under stress (e.g., exams). Then, they will be more confident in how they should answer the question, and which aspects are worth focusing on.

We can sum up the process of formulating the criteria for the assessment of a problem with the help of **Figure 4**. It shows these processes as two autonomous systems. The first is the learning environment in which the students discuss the criteria. The second is based on the teacher's observation of what was happening in the group, the products, and the subsequent reflection. As a result of this, there is an opportunity for the teacher to learn about the students' perceptions of the components of the problem and reflect upon the quality of the teaching that occurred in the classroom.

Overall Benefits to Students and Teacher

An activity of this kind, in which students negotiate with each other to formulate criteria, has been shown to have the potential of creating greater clarity concerning the necessary elements required in the solution of a mathematical problem, which points to a higher quality of work (cf. Topping, 2003). This finding also emerged in this study: the students reported that having to formulate the criteria for solving a problem was a very valuable task that enables them to better understand the underlying properties of the question. The teacher's remarks reinforced this, as she stated that she believed that engaging students in thinking about the components of the question and not only about the final product is a sound strategy for solving similar problems at a later stage.

Both the students and teacher noted that this process was allowed coordinating with each other's expectations. Also, as Sadler (1989) observed, teachers sometimes have difficulty expressing to themselves and others what they know. By observing students in the peer assessment activity and by discussing their actions with them, teachers are better able to construe what their students have understood from the lessons and therefore how successful their lessons were in relaying the concepts. This happened to the teacher in our study who, after reviewing the video documentation, realized that she had omitted the concepts of 'limited domain' and 'function's domain' in her lessons and concluded that she would have to deal with those matters more explicitly in her lessons.

CONCLUSION

Recent empirical findings highlight the dual nature of peer assessment in mathematics education. On one hand, students report that engaging in peer evaluation enhances their conceptual understanding, metacognitive awareness, and appreciation for multiple problem-solving strategies. Biton (2025a, 2025b) demonstrated that students valued the opportunity to construct assessment criteria and reflect on their peers' mathematical thinking, noting increased motivation and collaborative engagement. On the other hand, Biton's (2025a, 2025b) study also showed that the students raised significant concerns regarding fairness, reliability, and the influence of social dynamics on evaluative judgment and underscored students' hesitations about the legitimacy of peer-generated scores within formal grading frameworks, especially when anonymity and structured training were lacking.

These findings point to the need for careful instructional design that incorporates rater preparation, collaborative norm-setting, and, where appropriate, digital tools to foster objectivity and reduce interpersonal biases. This suggests that peer assessment, when appropriately scaffolded, holds promise as a formative tool that supports critical thinking, mathematical reasoning, and learner agency.

The present study suggests that the process of peer assessment led to meaningful interaction among the student assessors and a profound enhanced understanding of the concepts that were involved in the mathematical problems. Despite any heterogeneity in a group, we noted that all the students actively participated in negotiating evaluation criteria, the mathematical validity of the solutions being evaluated, and the value of each of the solution components in the grading processes. In addition, having to consciously decide on criteria and scoring led students to better understand their teachers, and observing the actions of the students helped the teacher understand what aspects were missing in her lessons. Our hope is that future research will divulge new relevant insights which we consider important to the future of mathematics teaching.

Ethical statement: The author stated that the study was conducted in accordance with applicable ethical research standards, and that participants provided written informed consent after receiving an explanation of the study's objectives and procedures.

Declaration of interest: No conflict of interest is declared by the author.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the author.

REFERENCES

Balacheff, N. (1991). The benefits and limits of social interaction: The case of mathematical proof. In A. Bishop, S. Mellin-Olson, & J. van Doormolen (Eds.), *Mathematical knowledge: Its growth through teaching* (pp. 175-192). Kluwer Academic. https://doi.org/10.1007/978-94-017-2195-0_9

Biton, Y. (2025a). Student reflections on peer assessments: Benefits and challenges in a mathematics class. *Educational Process: International Journal, 14*, Article e2025003. https://doi.org/10.22521/edupij.2025.14.3

Biton, Y. (2025b). Learning mathematics through peer assessment: "How can we assess something that we ourselves don't know how to solve?" *Eurasia Journal of Mathematics, Science and Technology Education, 21*(1), Article em2557. https://doi.org/10.29333/ejmste/15794

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

Boud, D. (1990). Assessment and the promotion of academic values. *Studies in Higher Education, 15*(1), 101-111. https://doi.org/10.1080/03075079012331377621

Boud, D., Cohen, R., & Sampson, J. (1999) Peer learning and assessment. *Assessment & Evaluation in Higher Education, 24*(4), 413-426. https://doi.org/10.1080/0260293990240405

Brown, S., & Knight, P. (1995). *Assessing learners in higher education*. Kogan Page.

Cartney, P. (2010). Exploring the use of peer assessment as a vehicle for closing the gap between feedback given and feedback used. *Assessment & Evaluation in Higher Education, 35*(5), 551-564. https://doi.org/10.1080/02602931003632381

Cho, K., & MacArthur, C. (2010). Student revision with peer and expert reviewing. *Learning and Instruction, 20*(4), 328-338. https://doi.org/10.1016/j.learninstruc.2009.08.006

Davidson, G. V. (1990). Matching learning styles with teaching styles: Is it a useful concept in education? *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), Article 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

Falchikov, N. (2001). *Learning together: Peer tutoring in higher education*. Routledge .

Gielen, S., Peeters, E., Dochy, F., Onghena, P., & Struyven, K. (2010). Improving the effectiveness of peer feedback for learning. *Learning and Instruction, 20*(4), 304-315. https://doi.org/10.1016/j.learninstruc.2009.08.007

Harel, G., & Koichu, B., (2010). An operational definition of learning. *Journal of Mathematical Behavior, 29*, 115-124. https://doi.org/10.1016/j.jmathb.2010.06.002

Johnson, R. T., & Johnson, D. W. (1985). Student-student interaction: Ignored but powerful. *Journal of Teacher Education, 34*(36), 22-26. https://doi.org/10.1177/002248718503600406

Keppell, M., & Carless, D. (2006). Learning-oriented assessment: A technology-based case study. *Assessment in Education, 13*(2), 179-191. https://doi.org/10.1080/09695940600703944

Keppell, M., Au, E., Ma, A., & Chan, C. (2006). Peer learning and learning-oriented assessment in technology-enhanced environments. *Assessment & Evaluation in Higher Education, 31*(4), 453-464. https://doi.org/10.1080/02602930600679159

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

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

Leikin, R., & Zaslavsky, O. (1999). Connecting research with practice: Cooperative learning in mathematics. *Mathematics Teacher, 92*, 240-246. https://doi.org/10.5951/MT.92.3.0240

McLaughlin, P., & Simpson, N. (2004). Peer assessment in first year university: How the students feel. *Studies in Educational Evaluation, 30*, 135-149. https://doi.org/10.1016/j.stueduc.2004.06.003

Moschkovich, J. N. (1996). Moving up and getting steeper: Negotiating shared descriptions of linear graphs. *The Journal of the Learning Sciences, 5*(3), 239-277. https://doi.org/10.1207/s15327809jls0503_4

Moschkovich, J. N. (1998). Resources for refining conceptions: Case studies in the domain of linear functions. *The Journal of the Learning Sciences, 7*(2), 209-237. https://doi.org/10.1207/s15327809jls0702_3

NCTM. (2009). *Guiding principles for mathematics curriculum and assessment*. National Council of Teachers of Mathematics.

Norton, A., McCloskey, A., & Hudson, R. (2011). Prediction assessments: Using video-based predictions to assess prospective teachers' knowledge of students' mathematical thinking. *Journal of Mathematics Teacher Education, 14*, 305-325. https://doi.org/10.1007/s10857-011-9181-0

Orsmond, P., Merry, S., & Reiling, K. (1996). The importance of marking criteria in the use of peer assessment. *Assessment & Evaluation in Higher Education, 21*, 239-250. https://doi.org/10.1080/0260293960210304

Orsmond, P., Merry, S., & Reiling, K. (1997a). A study in self-assessment: Tutor and students' perceptions of performance criteria. *Assessment & Evaluation in Higher Education, 22*(4), 357-369. https://doi.org/10.1080/0260293970220401

Orsmond, P., Merry, S., & Reiling, K. (1997b). Students' and tutors' perceptions of a good essay. *Research in Education 58*, 81-83. https://doi.org/10.1177/003452379705800111

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*(1), 23-38. https://doi.org/10.1080/02602930050025006

Piaget, J. (1977). *The development of thought: Equilibration of cognitive structures*. Viking.

Radford, L. (2011). Classroom interaction: Why is it good, really? *Educational Studies in Mathematics, 76*, 101-115. https://doi.org/10.1007/s10649-010-9271-4

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*, 119-144. https://doi.org/10.1007/BF00117714

Sambell, K., McDowell, L., & Sambell, A. (2006). Supporting diverse students: Developing learner autonomy via assessment. In C. Bryan, & K. Clegg

(Eds.), *Innovative assessment in higher education* (pp. 158-168). Routledge.

Saxe, G. B., Gearhart, M., Shaughnessy, M., Earnest, D., Cremer, S., Sitabkhan, Y., Platas, L., & Young, A. (2009). A methodological framework and empirical techniques for studying the travel of ideas in classroom communities. In B. Schwartz, T. Dreyfus, & R. Hershkowitz (Eds.), *Transformation of knowledge in classroom interaction* (pp. 203-222). Routledge.

Somervell, H. (1993). Issues in assessment, enterprise and higher education: The case for self-, peer and collaborative assessment. *Assessment & Evaluation in Higher Education, 18*, 221-233. https://doi.org/10.1080/0260293930180306

Strachan, I. B., & Wilcox, S. (1996). Peer and self assessment of group work: Developing an effective response to increased enrolment 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: Unravelling peer assessment: methodological, functional, and conceptual developments. *Learning and Instruction, 20*, 265-269. https://doi.org/10.1016/j.learninstruc.2009.08.002

Taylor, S., Rizvi, F., Lingard, B. & Henry, M. (1997). *Educational policy and the politics of change*. Routledge.

Topping, K. J. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research, 68*(3), 249-276. https://doi.org/10.3102/00346543068003249

Topping, K. J. (2003). Self and peer assessment in school and university: Reliability, validity and utility. In M. Segers, F. Dochy, & E. Cascallar (Eds.), *Optimising new modes of assessment: In search of qualities and standards* (pp. 55-87). Kluwer Academic Publishers. https://doi.org/10.1007/0-306-48125-1_4

Topping, K. J., Li, H., Xiong, Y., Hunter, C. V., Guo, X., & Tywoniw, R. (2020). 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

Webb, N. (1991). Task-related verbal interaction and mathematics learning in small groups. *Journal for Research in Mathematics Education, 22*, 366-389. https://doi.org/10.5951/jresematheduc.22.5.0366

Williams, E. (1992). Student attitudes towards approaches to learning and assessment. *Assessment & Evaluation in Higher Education, 17*, 45-58. https://doi.org/10.1080/0260293920170105

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*. 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: Proceedings of the 5th Annual Teaching Learning Forum* (pp. 189-200). Murdoch University.