

RESEARCH

Open Access



Self vs. peer assessment activities in EFL-speaking classes: impacts on students' self-regulated learning, critical thinking, and problem-solving skills

Tribhuwan Kumar¹, Seyyed Mohammad Ali Soozandehfar², Arash Hashemifardnia^{3*} and Rasoul Mombeini⁴

*Correspondence:
arashhashemi1184@gmail.com

¹ Department of English Language and Literature, College of Science and Humanities at Sulail, Prince Sattam Bin Abdulaziz University, Al Kharj, Saudi Arabia

² Department of TEFL, University of Hormozgan, Bandar Abbas, Iran

³ Department of English, Shahrekord Branch, Islamic Azad University, Shahrekord, Iran

⁴ English Department, Ahvaz Branch, Islamic Azad University, Ahvaz, Iran

Abstract

Assessment has several advantages like involving students in learning and making them cognizant of their strengths and weaknesses. The effects of peer assessment and self-assessment as two main kinds of assessment have not examined on EFL learners' self-regulated learning, critical thinking, and problem-solving skills; therefore, the present research attempted to inspect the impacts of self and peer assessment activities on self-regulated learning, critical thinking, and problem-solving skills of Iranian EFL learners. To do so, 75 Iranian EFL learners were selected and divided into two experimental groups (peer assessment group (PAG) and self-assessment group (SAG)) and one control group (CG). Then, three questionnaires of self-regulated learning, critical thinking, and problem-solving skills were administered to the students as the pre-tests of the study. After that, one experimental group received the treatment via using self-assessment and the other experimental group received the instruction through peer assessment. The control group received a teacher-based assessment instruction. After teaching 13 English conversations to all groups, 3 post-tests were given to them to measure the impacts of the treatment on their self-regulated learning, critical thinking, and problem-solving skills. The results of one-way ANOVA showed that there was a substantial difference between the post-tests of the experimental groups and the control group. The results showed that both experimental groups outperformed the control group in the post-tests of self-regulated learning, critical thinking, and problem-solving skills. Furthermore, the outcomes revealed that there were no significant differences between the post-tests of the experimental groups. It can be concluded that both types of assessments play a vital role in English language teaching and learning. Based on the results of this study, language teachers, materials developers, and education authorities can concentrate on using self and peer assessments as one of the main strategies to develop the language learning ability of Iranian EFL learners.

Keywords: Critical thinking, Peer assessment, Problem-solving skills, Self-assessment, Self-regulated learning

Introduction

When working within the alternative assessment framework, self-assessment and peer assessment are recognized as two favorite and highly recommended practices (Brown and Hudson, 2012). According to Birjandi and Bolghari (2015), self-assessment is related to the involvement of the learners in measuring their learning, mainly their achievements and their learning outcomes. Cooperative learning and peer assessment are related (Fulcher and Davidson, 2007; Tigchelaar, 2019). This strategy emphasizes the value of collaboration and encourages more student contact. It is also one of the choices for combining instruction and evaluation. According to research, the guiding principles of self- and peer assessment have some benefits (Brown and Abeywickrama, 2010).

A counterargument to criticism of conventional systems of assessment is self-assessment. Self-assessment is considered one of various assessments which may be executed by the students in exploring, growing, and figuring out their functionality in regards to the course (Marzuki et al. 2020). The practice of accepting accountability for assessment and learning allows students to acquire more independence and self-management (Chen, 2008). According to Birjandi and Bolghari (2015), self-assessment is linked to learners' participation in assessing their learning, namely their instructional consequences, and accomplishments. The application of self-assessment in the classes can inspire pupils to have an active participation in learning. To inspire learners to think critically about their learning outcomes and process, this style of assessment is usually utilized for formative assessment purposes (Harris and Brown, 2013). Additionally, Musfirah (2019) contends that self-assessment is advantageous since it is a better approach for students straightly about their problems.

Richards and Schmidt (2002) define peer evaluation as "an alternative assessment strategy in which learners assess the progress of each other by applying some checklists supplied by their instructors" (p. 47). It works well as a substitute for traditional assessment. Additionally, according to Saito and Fujita (2009), it is one of the approaches that can integrate training and assessment.

Peer assessment is thought to be one of the primary categories of option evaluation. The need for associate evaluation is emphasized by both differentiated informative learning and instructional examination (Hung, 2019). According to Slavin (1997), the associate appraisal was one of the greatest achievements in educational history. Peer review enhances students' learning by fostering "a sense of proprietorship and obligation, inspiration, and impression of the understudies' own learning" (Saito and Fujita, 2009, p. 152). Peer assessment is one of the useful techniques for evaluating students in the classroom, according to Shepard (2000).

According to Brown and Hudson (2012), the guiding concepts of self- and peer assessment offer a number of advantages. One of the newer forms of assessment that language learners use to explore, improve, and comprehend their functionality is self-assessment (Ma and Winke, 2019). Self-assessment has been touted as an effective evaluation method for language learners due to its critical role in maintaining students' attention, engagement, motivation, accountability, and self-direction (Miller, 2003). Peer evaluation, on the other hand, is thought to enhance language learners' self-confidence, accountability, sense of ownership, social skills, and ability to work cooperatively and collaboratively (Esfandiari and Tavassoli, 2019). Appreciating learners' internal

characteristics and how these aspects may affect learning and evaluation is one of the fundamental tenets of using alternative assessment (McKay, 2006).

Using self- and peer assessments can develop EFL learners' self-regulation. Based on Pedrotti and Nistor (2019), self-regulation is the process through which learners try to direct their own learning by setting goals, planning how to achieve them, evaluating the learning process, applying problem-solving learning approaches, and gauging their own progress. The ability to organize actions and the learning process come next after the development of skills and reasoning through independent learning (Kuiper-Anne and Pesut, 2016).

According to Schunk and Zimmerman (2012), the independent learning process is characterized by students taking proactive steps to ensure the continuity of their education by using their knowledge, formulating plans, modifying current effects, boosting their learning confidence, and making decisions. According to the assertion, Hadwin et al. (2015) found that independent learning, in which students are active agents who can manage their learning, affects rising interest and willingness in learning. Students are more effective when they independently participate in behaviors that govern their desire to learn, such as planning what to learn and reviewing the subject matter, according to research by Wong et al. (2021). This is consistent with Clark's (2012) claim that studying on one's own improves motivation and academic performance since students can adaptively comprehend the learning qualities in accordance with their capacities. Hadwin et al. (2015) later provided an explanation for this, stating that imitation and the formation of cognitive patterns and strategies correspond to how it appears to follow observational learning (modeling).

The other variable that can be affected by self- and peer assessments is problem-solving learning. According to Argaw et al. (2017), problem-solving is a learning strategy used to give context and motivation for solving an issue. The formation of students' problem-solving skills, according to Chua, Tan, and Liu (2016), is built on four key stages: issue-solving, problem analysis, discovery and reporting, and evaluation to find answers. Han and Toh (2019) emphasize that the progress of students' skills and criticism in investigating any knowledge has been influenced by their motivation to solve difficulties (Chua et al. 2016; Rezai et al. 2022; 2023). Hu et al. (2017) discovered that a lot of educational personnel employed problem-solving techniques to get around challenges with scientific learning. According to Laurens et al. (2018), problem-solving in science offers solutions for resolving everyday issues that serve as the foundation for actions and subsequent steps. This is consistent with the assertion made by Sukariasih et al. (2020), who stress that problem-solving in physics learning helps students build skills to tackle difficulties in the real world. Fitriani et al. (2020) identify problem-solving as a technique that affects how knowledge and ideas are built.

Applying self- and peer assessments can also enhance the critical thinking of EFL learners. Critical thinking is viewed as a sort of critical analysis, and according to Bobkina and Stefanova (2016), it is "trained intellectual criticism that combines study, understanding of historical context, and balanced judgment: It is the capacity for rational and analytical thought. According to Paul and Elder (2008), critical thinking is the deliberate and reflective decision of what to believe or what to do in response to experience, observation, verbal or written expressions, or arguments. Deciding the relevance and meaning

of what is observed or articulated or, in the case of an inference or argument, deciding if there is sufficient support for accepting the conclusion as true, are all examples of critical thinking.

The definition provided by Otite and Ogionwo (2006), which is “active and skilled interpretation and evaluation of information, communication, observation, and argumentation,” is consistent with this description. Therefore, in critical thinking, the evidence, the context of the judgment, the pertinent standards for making the judgment well, the appropriate methods or techniques for forming the judgment, and the appropriate theoretical framework for comprehending the issue and the question at hand are all given due consideration. Broad intellectual standards including clarity, credibility, accuracy, precision, depth, relevance, breadth, fairness, and significance are also used in critical thinking in addition to logic. The word “critical” has a negative connotation in modern usage, which is not always true with critical thinking. For instance, a critical analysis of an argument can conclude that it is persuasive (Thomas and Nelson Laird, 2010).

The defined variables play a significant role in language teaching and learning. They produce positive effects both for EFL teachers and learners. Many EFL teachers believe that assessment is a vital component to improve the student’s abilities to evaluate their own performance leading to its improvement. Concerning the significance of assessment, the current study tried to examine the effects of self and peer assessment activities on Iranian EFL learners’ self-regulated learning, critical thinking, and problem-solving skills. Doing this research can be significant as it deals with some important psychological variables that have direct impact on language learning. Also, it is important as it can bring about some useful implications for EFL teachers and learners. Peer and self-assessment can inspire students to take greater responsibility for their learning as they assess each other and themselves. Self-assessment makes students reflect on their own work and judge how well they have performed in relation to the assessment criteria. Also, it provides some opportunities to be able to identify what constitutes a good piece of work. Peer assessment encourages student responsibility and involvement; encourages students to reflect on their role and contribution to the process of group work; focuses on the development of student’s judgment skills; and provides more relevant feedback to students as it is generated by their peers.

Literature review

Theoretical background

Appropriate evaluation is one of several aspects that can help EFL students’ reading comprehension. Regardless of the instructional strategies employed, assessment is crucial to the teaching and learning processes since it gives students feedback (Alias et al. 2015). Due to its emphasis on using real-world contexts, identifying learners’ strengths and weaknesses, fostering human judgment, and utilizing open disclosure of principles and rating standards, assessment can actually be applied as a useful strategy to develop learning various facets of a foreign or second language (Ashraf and Mahdinezhad, 2015; Namaziandost et al. 2023). Being active in assessment processes, procedures, and results is important for instructors and students alike claim Cheng and Warren (2005). Utilizing

self-assessment is a popular technique of evaluation to get students engaged in assessment and accountable for their own learning.

For the classroom teacher, self-evaluation is an appropriate alternative to traditional techniques of assessment. It is a specific metacognitive strategy that demands particular attention. Additionally, it assists in the growth of the qualities of “good language learners, which involves the ability to assess their own performance and the ability to be self-critical” (Hedge, 2000, p.94). By putting more emphasis on enabling pupils to set their own objectives and continue their development, it also aids in the progress of learners’ self-governing learning abilities (Butler and Lee, 2010; Ghahderijani et al. 2021). Self-assessment proponents oppose that trusting only on teacher evaluation results in students not taking ownership of their own learning and support depending on the teacher (Hung, 2018).

According to Boud (2000), to become effective learners throughout their lives, students must learn how to assess their own performance to get autonomous of their teachers. Similarly, Miller (2003) thought that self-assessment was a beneficial strategy for students because it plays a big part in keeping students engaged, responsible, and autonomous. The need for learners to become familiar with the standard scoring similar to their teacher’s and the proper method of evaluating their own performances is a crucial concern in the self-assessment process (Babaii et al. 2016).

One of the main types of alternative assessment, known as “peer assessment,” has drawn a lot of interest in instructional learning and educational research. It is regarded as a system in which people take into account the quantity, degree, worth, value, or success of the outputs or results of the learning of peers with comparable standing (Topping, 2009). It involves “the act of getting readers to evaluate what they’ve read, and perhaps offer suggestions for how their peers should be graded” (Roberts, 2006, p. 80), as well as being evaluated on the caliber of the evaluations made (Davies, 2006). Considerable advantages of peer evaluation include immediate help in the classroom, gains for both the assessors and the assessees, and being individualized and engaging (Black and William, 1998; Fu et al. 2019).

Peer evaluation, according to Saito (2008), promotes reflective learning by fostering awareness of performance standards through observation of others’ performances. Although some students have fears and misgivings, in general, peer assessment seems to cause favorable emotions in learners. It promotes the development of self-awareness by highlighting the discrepancy between one’s perception and that of others, allowing for further learning and accountability for it. Concentrating on peers’ weaknesses and strengths can also improve pupils’ learning, develop their critical thinking skills, and help them become more autonomous. Zhi-Feng and Yi Lee (2013) claim that after taking part in peer assessment activities, the students improved their work by using feedback from others. The majority of pupils felt well about peer observation.

Teachers initiate discussions about assessment with students based on the most recent advancements in learning theories; this is a significant problem for assessment in the twenty-first century since it places demands on the teachers to acquire the specialized skills required for this new, extra role. More intensive, interactive methods should be utilized to evaluate the learning process, and this work should be done in groups of

peers or between a teacher and student (Adachi et al. 2018; Azizi and Namaziandost, 2023; Matsuno, 2009).

In conclusion, the options for language assessment are peer and self-assessment. Based on Falchikov (2005, p. 27), “learners apply standards and use criteria to the work of their peers to evaluate that work” during peer assessment. Building on the latter, students examine their own work using standards and criteria in self-assessment. Peer and self-evaluation are anticipated to lessen the teacher’s dominant role in assessment activities. Peer and self-assessment have become more extensively utilized in higher educational contexts during the past few decades (Cheng and Warren, 2005; Tavakoli, 2010; Weisi and Karimi, 2013). In peer assessment, some support and scaffolding are provided by the classmates and teachers which they are rooted in Vygotsky’s (1978) idea of the Zone of Proximal Development (ZPD). These assistance and support do not come from an authoritative teacher but from a competent peer, partner, or classmate. ZPD is described as the distance between an individual’s independent functioning performance and the potential performance level with others’ assistance.

Empirical background

With an emphasis on speaking ability, Ashraf and Mahdinezhad (2015) explored the influence of self- and peer assessment in fostering autonomy in language use. Forty-eight EFL Iranian students with a comparable level of proficiency were selected among 110 students for the study and divided into two groups: the control group (CG) and the experimental group (EG). They were then given pre-tests for their speaking skills and learner autonomy. Throughout 12 sessions, the students in the experiment students were required to assess the language of their classmates using a variety of exercises. The learners in the control class were asked to evaluate their own language using the same standard teaching and learning procedures. The same assessments were given to both student groups as a final test. The examination of the data revealed that respondents in the experimental group did better than those in the control group, demonstrating the beneficial impact of peer evaluation on learners’ autonomy. Additionally, the analysis of the speaking test data revealed notable differences between the experimental and control students. It can be said that self-evaluation has less of an impact on the autonomy and speaking ability of EFL learners than peer assessment.

Khairani (2019) examined how grade 10 students in Indonesia used the PAs in task-based learning to improve their reading skills. Four processes were used in this study’s qualitative research methodology: data collection, data reduction, data presentation, and conclusion and verification. Three English instructors of pupils in grade tenth provided the study’s data, which were drawn from PA used by the teachers. Performing documentation, conducting interviews, and making observations were the methods used to collect the data. The study’s findings showed that two teachers effectively used PAs, but one instructor did not. The results showed that task-based learning was a successful method for using PAs in RC. In a different study, Seifert and Feliks (2019) looked at the impacts of PA and self-assessment to ensure that students would take more ownership of their learning. Their findings showed that the participants benefited significantly from these two assessments to advance their learning and assessment skills.

Meletiadou (2021) used a pre-test post-test quasi-experimental method to investigate the influence of PA on the writing performance of 200 Greek Cypriot EFL students. For the whole school year, these adolescent students attended two writing lessons every week (90 min). Teachers were trained in PA skills before having to train their own students. Students were instructed to utilize a PA rubric developed by the researcher but agreed between the students and their teachers throughout the training sessions. Paired *t* tests were used to see if students in the control ($n = 100$ students and 10 teachers) and experimental ($n = 100$ students and 10 teachers) groups improved their writing performance when comparing pre- to post-test scores. The study outcomes indicated that PA could have a moderately positive impact on students' writing performance. The use of PA improved students' writing performance in 5 aspects: mechanics, organization, content, focus, and vocabulary and language use.

Tunagür (2021) conducted a quantitative study in Turkey to determine whether or not employing PAs had a favorable impact on sixth-graders motivation for writing and writing anxiety. Thirty-five students participated in this study: 17 were in the EG group and 18 were in the CG group. A writing anxiety questionnaire and a writing motivation questionnaire were used to collect the research data. The PA was employed throughout 40 days, and the texts that the participants wrote were evaluated by their peers using the PA forms. First, the *t* tests for unrelated groups and two-way ANOVA testing for complex measures were employed to analyze the collected data. Based on the study's findings, it was determined that the EG experienced significantly less writing anxiety than the CG. The results also revealed that the EG's motivation to write improved following treatment. According to the results, it can be said that applying PAs helps students enhance their writing motivation while reducing their writing anxiety.

In another study, Ritonga et al. (2022) looked at how peer evaluation affected the growth of Iranian EFL learners' reading motivation (RM), vocabulary learning (VL), and reading comprehension (RC). To accomplish this, 60 Iranian EFL learners were chosen and separated into two groups: the CG and EG. The groups then completed the RC, RM, and VL pre-tests. The EG was then split into six groups of five, and PA was used to assess how well they did in reading and vocabulary. The EG pupils evaluated their classmates' work with the help of the teacher. The students in the CG class evaluated their performance after each test with the help of the teacher. Both groups received the post-tests for vocabulary, RM, and RC after the 15-session treatment program. The one-way ANCOVA test findings revealed that the EG outperformed the CG on the three post-tests for RC, RM, and VL. In fact, the results demonstrated that PA positively affected the RC, RM, and VL of Iranian EFL students.

Hong Canh (2022) investigated the impacts of PA on EFL students' writing performance, focusing on lower secondary schools. One hundred students from Le Ngoc Han Lower Secondary School were divided into two groups: control and experimental. Students in the experimental group were given PA, while those in the control group were taught the standard method. To assess the student's writing performance before and after treatment, pre-test and post-test research instruments were used. A questionnaire was employed to assess students' views regarding PA in their learning writing at the same time. The study's findings revealed that PA significantly improves EFL students' writing abilities, and a substantial percentage of participants had good views towards PA in their writing learning. Therefore,

the teachers' implementation of PA in teaching writing in lower secondary schools should be encouraged with a view to enhancing the students' writing performance.

Imani (2022) compared how peer assessment and self-evaluation influenced impulsive and reflective EFL students' speaking ability. In order to do this, a set of 51 EFL learners (27 reflective and 24 impulsive) were selected and assigned to two experimental classes at random—peer assessment and self-assessment. The evaluation in the self-assessment group was based on Speaking Self-Assessment Sheets. The evaluation in the peer assessment group was based on a Peer Assessment Rating Sheet. Following the completion of the treatment, the individuals were given the speaking component of another PET as a posttest. The two-way ANOVA results revealed that (a) there was no significant interaction between assessment type and cognitive type, (b) peer assessment and self-assessment had the same impact on impulsive and reflective learners' speaking ability, and (c) regardless of assessment type, learners with different cognitive types performed differently, with reflective learners outperforming their impulsive counterparts.

Reviewing the literature indicates that using self- and peer assessments can be an effective tool for developing English language learning. Most studies in the domain of peer assessment and self-assessment were conducted on language skills (writing, speaking, listening, and reading and sub-skills (grammar, pronunciation, etc.)). No study was found to examine the effectiveness of the mentioned assessments on EFL learners' self-regulated learning, critical thinking, and problem-solving skills. Therefore, the current research intended to inspect the effects of self- and peer assessment on enhancing Iranian EFL learners' self-regulated learning, critical thinking, and problem-solving skills. Consequently, the following research question was raised:

RQ1. Does the use of self- and peer assessments have any significant effect on EFL learners' self-regulated learning?

RQ2. Does the use of self- and peer assessments have any significant effect on EFL learners' critical thinking?

RQ3. Does the use of self- and peer assessments have any significant effect on EFL learners' problem-solving skills?

Regarding the research questions, three null hypotheses were formulated in this study:

HO1. The use of self- and peer assessments does not have any significant effect on EFL learners' self-regulated learning.

HO2. The use of self- and peer assessments does not have any significant effect on EFL learners' critical thinking.

HO3. The use of self- and peer assessments does not have any significant effect on EFL learners' problem-solving skills.

Method

Research design

A quantitative method and a quasi-experimental design were used by the researchers to collect the needed data. In fact, we used pre-test-treatment-post-test design for

gathering the data. Two experimental groups and one control group were included in our study. Self-assessment and peer assessment are the independent variables and our dependent variables are self-regulated learning, critical thinking, and problem-solving skills.

Participants

The Oxford Quick Placement Test (OQPT) results determined the selection of 75 participants from a panel of 99 students. They were picked from Ahvaz, Iran's Mahan English Language Institute. The respondents' English skill level was intermediate, and their ages ranged from 21 to 33. The participants' first language was Persian, and they were all men. This study contained two equal experimental groups of PAG and SAG, as well as a CG.

Instruments

The first instrument utilized in this research was the OQPT which was regarded as the proficiency test to determine the degree of competency of the participants. The OQPT comprises 200 tests that assess hearing, grammar, vocabulary, and reading skills. OPT is a dependable and effective method of placing students at various levels of language competence. Allen (2004), the test's creator, argues that the OQPT can be used with any number of English learners to assure rapid, reliable, and accurate grading and placement of students in classes ranging from elementary to advanced. The OQPT has been calibrated against competency levels based on the Common European Framework of Reference for Languages (CEFR), the Cambridge ESOL Examinations, and other important international examinations, according to Allen. In the current study, the test's reliability was 0.81 as determined by Cronbach's alpha.

The second instrument of this study was a Self-Regulated Learning (SRL) questionnaire. Seker (2015) created and validated a five-point Likert-type questionnaire. It was used to determine learners' self-reported SRL utilization, which was derived from earlier SRL models. The questionnaire was mostly based on the Self-Regulated Learning Model developed by Boekaerts and Corno (2005) and the Strategy Inventory for Language Learning (SILL) developed by Oxford (1990). The questionnaire contained 30 items, which were divided into five subscales: internal motivation ($n = 5$), external motivation ($n = 4$), cognitive strategies ($n = 7$), metacognitive strategies ($n = 10$), and evaluation ($n = 4$). In the current study, the calculated Cronbach's alpha value was 0.81 at the pretest and 0.83 at the posttest.

The third tool applied in this research was the Watson-Glaser Critical Thinking Appraisal-Form A (1980). It is divided into five sections, each with 80 items that assess the five aspects of CT as defined by Watson and Glaser (1991): inference, recognizing assumptions, drawing deductions, interpretation, and evaluation. Watson and Glaser (2002) reported test-retest reliability ($r = .81$) in addition to the satisfactory face, content, construct, and criterion validity of this instrument. In the current study, Cronbach's alpha reliability of the Watson-Glaser Critical Thinking Appraisal (Form A) was .87.

The last instrument was a scale for measuring the Problem-Solving Skills (PSSs) of the participants. The assessment of student work outcomes in solving basic physics questions on mechanics material using grading rubrics ranging from 0 (zero) to 4 (four)

yielded data on student PSSs. A 0 (zero) score indicates that pupils are unable to create or use the indicator component. A student with a score of 1 (one) was able to construct or use the indicator component, however, there were some errors. Students with a score of 2 (two) were able to produce or use the indicator component, but it was incomplete and imprecise. Students with a score of 3 (three) were able to make or use the indication component entirely and accurately. Students with a score of 4 (four) were able to build or apply the indication component extremely completely and precisely. Docktor and Heller (2009) created five indications on the Robust Assessment Instrument for Student Problem-solving, namely visualization/problem description, physics approach, a unique application of physics principles, mathematical processes, and logical conclusions. The SRL and PSS data were analyzed descriptively. The hypothesis was then tested using the Partial Least Square (PLS) analysis technique with the Smart PLS 3.0 program to determine the influence of SRL on student PSS in online Basic Physics courses. Four phases are utilized to verify the validity and reliability of the reflective measurement model in PLS analysis: internal consistency reliability, convergent validity, indicator reliability, and discriminant validity of the new model. The calculated Cronbach's alpha value was 0.84 for the PSSs questionnaire. The validities of all instruments were confirmed by three English professors.

Data collection procedure

To conduct this study, the OQPT was initially administered to the respondents to measure their level of English proficiency. Seventy-five people were chosen from a pool of 99 to represent the study's target group. Following that, the participants were divided into two experimental groups (PAG and SAG) and a CG at random. Following that, all groups received the self-regulated learning, critical thinking, and problem-solving skills questionnaires as study pre-tests. Then, the members of one experimental group received the treatment based on the self-assessment instruction. With the teacher's help, the participants in the self-assessment group analyzed their own performances following each test and made comments about their weak and strong points. The other experimental group received the treatment based on the peer assessment instruction. This experimental group was divided into five sub-groups and peer assessment was used to assess their speaking performances. At the outset of each session, one test including some speaking items was given to this group and they were required to check their classmates' performances in peers with the guidance of the teacher. The students in the control group, on the other hand, received the treatment based on a conventional assessment; the teacher assessed the students' performances. After teaching 13 conversations to each group, the aforementioned questionnaires were used once more after the instruction to determine how the participants' self-regulated learning, critical thinking, and problem-solving skills were affected by the use of mentioned treatment.

Data analysis

To assess the gathered data, SPSS software, version 22, was exploited. The descriptive statistics were first calculated. After that, a one-way ANOVA test was performed to measure how the instruction affected the students' self-regulated learning, problem-solving skills, and critical thinking.

Table 1 Descriptive statistics of self-regulatory learning pre-test

	N	Mean	Std. deviation	Std. error	95% confidence interval for mean		Minimum	Maximum
					Lower bound	Upper bound		
CG	25	63.16	7.318	1.46	60.13	66.18	50.00	78.00
SAG	25	63.84	6.40	1.28	61.19	66.48	54.00	78.00
PAG	25	60.60	11.33	2.26	55.92	65.27	50.00	85.00
Total	75	62.53	8.62	.99	60.55	64.51	50.00	85.00

Table 2 Inferential statistics of self-regulatory learning pre-test

	Sum of squares	df	Mean square	F	Sig.
Between groups	145.94	2	72.97	.98	.38
Within groups	5352.72	72	74.34		
Total	5498.66	74			

Results

The researchers obtained the relevant data, which they then utilized to examine the results. The specifics of the findings are shown in the tables below.

Table 1 displays the descriptive statistics for the three groups. All groups’ means are nearly equal. The self-evaluation group had a mean score of 63.84, the peer assessment group had a mean score of 60.60, and the control group had a mean score of 63.16. This indicates that since the groups were homogeneous at the start of the treatment, they were all in some ways comparable.

Table 2’s one-way ANOVA test was used to determine whether there were any potentially significant differences between the three groups’ pre-test results. The difference in means between the sample groups is not statistically significant at ($p < 0.05$) because Sig. level (.38) is more than 0.05 as the yardstick to put means to the test. On the self-regulatory learning pre-test, both the experimental groups and the control group displayed equal performance.

The results of the self-regulatory learning post-test for the control group, the self-assessment group, and the peer assessment group are shown in Table 3 using descriptive statistics. In actuality, the control group, the self-assessment group, and the peer evaluation group each had mean scores of 65.40, 118.04, and 119.32. This suggests that on the self-regulatory learning post-test, the three aforementioned groups’ performances varied.

According to Table 4, the difference between the experimental and control groups is significant at ($p < 0.05$), with the strength of the Sig (.00) being less than 0.05. On the post-test, the experimental groups outperformed the control group.

The mean scores on the self-regulatory learning post-test for each group are contrasted in Table 5. According to the findings of the analysis of the data in the table above, there was a significant difference between the conditions, $P < 0.05$. This means that there is a significant difference ($p < 0.05$) between the post-test results of the two experimental groups and the results of the control group. This table also demonstrates

Table 3 Descriptive statistics of self-regulatory learning post-test

	N	Mean	Std. deviation	Std. error	95% confidence interval for mean		Minimum	Maximum
					Lower bound	Upper bound		
Control	25	65.40	6.82	1.36	62.58	68.21	52.00	79.00
Self	25	118.04	10.88	2.17	113.54	122.53	108.00	160.00
Peer	25	119.32	10.30	2.06	115.06	123.57	108.00	160.00
Total	75	100.92	26.97	3.11	94.71	107.12	52.00	160.00

that there is no statistically significant difference between the peer and self-assessment groups on the post-test for self-regulatory learning ($P > 0.05$).

The critical thinking pre-test descriptive statistics for each of the three groups are shown in Table 6. In fact, the control group, the self-assessment group, and the peer evaluation group each had mean scores of 117.48, 115.12, and 115.76. This suggests that the three means of each group performed roughly equally on the critical thinking pre-test.

Table 7 demonstrates that the difference between the groups is not significant at ($p < 0.05$) since Sig (.77) is bigger than 0.05. They did equally well on the pre-test for critical thinking.

Table 8 displays the descriptive post-test data for critical thinking for each group. The mean scores for the control group, the self-assessment group, and the peer evaluation group are 111.96, 141.80, and 143.88, respectively. This indicates that the groups performed differently on the critical thinking post-test.

According to Table 9, the difference between the groups is significant at ($p < 0.05$) since Sig (.00) is less than 0.05. In actuality, experimental groups fared better on the post-test of critical thinking than the control group.

Table 4 Inferential statistics of self-regulatory learning post-test

	Sum of squares	Df	Mean square	F	Sig.
Between Groups	47333.12	2	23666.56	261.81	.00
Within Groups	6508.40	72	90.39		
Total	53841.52	74			

Table 5 Bonferroni test (multiple comparison of self-regulatory learning post-test)

(I) groups	(J) groups	Mean difference (I-J)	Std. error	Sig.	95% confidence interval	
					Lower bound	Upper bound
Control	Self	- 52.64 ^a	2.68	.00	- 59.23	- 46.04
	Peer	- 53.92 ^a	2.68	.00	- 60.51	- 47.32
Self	Control	52.64 ^a	2.68	.00	46.04	59.23
	Peer	- 1.28	2.68	1.00	- 7.87	5.31
Peer	Control	53.92 ^a	2.68	.00	47.32	60.51
	Self	1.28	2.68	1.00	- 5.31	7.87

^aThe mean difference is significant at the 0.05 level

Table 6 Descriptive statistics of critical thinking pre-test

	N	Mean	Std. deviation	Std. error	95% confidence interval for mean		Minimum	Maximum
					Lower bound	Upper bound		
Control	25	117.48	12.17	2.43	112.45	122.50	91.00	136.00
Self	25	115.12	12.62	2.52	109.91	120.32	90.00	134.00
Peer	25	115.76	11.49	2.29	111.01	120.50	95.00	134.00
Total	75	116.12	11.98	1.38	113.36	118.87	90.00	136.00

Table 7 Inferential statistics of critical thinking pre-test

	Sum of squares	Df	Mean square	F	Sig.
Between Groups	74.48	2	37.24	.25	.77
Within Groups	10551.44	72	146.54		
Total	10625.92	74			

Table 8 Descriptive statistics of critical thinking post-test

	N	Mean	Std. deviation	Std. error	95% confidence interval for mean		Minimum	Maximum
					Lower bound	Upper bound		
Control	25	119.96	9.38	1.87	116.08	123.83	100.00	136.00
Self	25	141.80	33.09	6.618	128.14	155.45	108.00	227.00
Peer	25	143.88	36.83	7.36	128.67	159.08	111.00	230.00
Total	75	135.21	30.69	3.54	128.15	142.27	100.00	230.00

Table 9 Inferential statistics of critical thinking post-test

	Sum of squares	df	Mean square	F	Sig.
Between groups	8778.98	2	4389.49	5.18	.00
Within groups	60959.60	72	846.66		
Total	69738.58	74			

The mean scores on the post-test for critical thinking are compared in Table 10 for all groups. According to the aforementioned table, both the experimental groups and the control group’s post-test results differ significantly ($p < 0.05$). This table demonstrates that there is no appreciable difference between the self- and peer assessment groups on the critical thinking post-test.

Table 11 displays the descriptive statistics for the three groups. The mean scores for the control group, the self-assessment group, and the peer evaluation group are 110.24, 112.12, and 113.04 respectively. This indicates that since the groups were homogeneous at the start of the treatment, they were all in some ways comparable.

Table 10 Bonferroni test (multiple comparison of critical thinking post-test)

(I) groups	(J) groups	Mean difference (I-J)	Std. Error	Sig.	95% confidence interval	
					Lower bound	Upper bound
Control	Self	- 21.84	8.23	.02	- 42.01	- 1.66
	Peer	- 23.92	8.23	.01	- 44.09	- 3.74
Self	Control	21.84	8.23	.02	1.66	42.01
	Peer	- 2.08	8.23	1.00	- 22.25	18.09
Peer	Control	23.92	8.23	.01	3.74	44.09
	Self	2.08	8.23	1.00	- 18.09	22.25

Table 12 demonstrates that the difference between the groups is not significant at (p 0.05) since Sig (.74) is bigger than 0.05. On the pre-test for problem-solving skills, they did equally well.

Table 13 displays the descriptive post-test data for critical thinking for each group. The mean scores for the control group are 112.72, 125.80 for the self-evaluation group, and 128.52 for the peer assessment group. This indicates that on the post-test of problem-solving ability, the groups' results varied.

Tables 14 and 15 show that the difference between the groups is significant at (p 0.05) since Sig (.00) is less than 0.05. In fact, the experimental groups outperformed the control group on the post-test of problem-solving ability.

On the problem-solving skill post-test, the means of all groups are compared (Table 15). As a result, both the experimental groups and the control group's post-test results differ significantly (p 0.05). This table demonstrates that there was no statistically significant difference between the self- and peer assessment groups on the problem-solving skill post-test.

To cut a long story short, the findings show that both experimental groups outdid the control group in all post-tests. Also, the results demonstrate no significant difference was found between the post-tests of both experimental groups; meaning that

Table 11 Descriptive statistics of problem-solving skill pre-test

	N	Mean	Std. deviation	Std. error	95% confidence interval for mean		Minimum	Maximum
					Lower bound	Upper bound		
Control	25	110.24	12.12	2.42	105.23	115.24	80.00	134.00
self	25	112.12	11.87	2.37	107.21	117.02	85.00	134.00
peer	25	113.04	14.76	2.95	106.94	119.13	80.00	150.00
Total	75	111.80	12.86	1.48	108.84	114.76	80.00	150.00

Table 12 Inferential statistics of problem-solving skill pre-test

	Sum of squares	df	Mean square	F	Sig.
Between groups	101.84	2	50.92	.30	.74
Within groups	12146.16	72	168.69		
Total	12248.00	74			

Table 13 Descriptive statistics of problem-solving skill post-test

	N	Mean	Std. deviation	Std. error	95% confidence interval for mean		Minimum	Maximum
					Lower bound	Upper bound		
Control	25	112.72	10.36	2.07	108.44	116.99	99.00	134.00
Self	25	125.80	13.29	2.65	120.31	131.28	99.00	140.00
Peer	25	128.52	13.42	2.68	122.97	134.06	100.00	145.00
Total	75	122.34	14.10	1.62	119.10	125.59	99.00	145.00

Table 14 Inferential statistics of problem-solving skill post-test

	Sum of Squares	df	Mean square	F	Sig.
Between groups	3567.70	2	1783.85	11.52	.00
Within groups	11147.28	72	154.82		
Total	14714.98	74			

Table 15 Bonferroni test (multiple comparison of problem-solving skill post-test)

(I) groups	(J) groups	Mean difference (I-J)	Std. error	Sig.	95% confidence interval	
					Lower bound	Upper bound
Control	Self	- 13.08	3.51	.00	- 21.70	- 4.45
	Peer	- 15.80	3.51	.00	- 24.42	- 7.17
Self	Control	13.08	3.51	.00	4.45	21.70
	Peer	- 2.72	3.51	1.00	- 11.34	5.90
Peer	Control	15.80	3.51	.00	7.17	24.42
	Self	2.72	3.51	1.00	- 5.90	11.34

self-assessment and peer assessment had equal effects on developing Iranian EFL learners' self-regulated learning, critical thinking, and problem-solving skills.

Discussion

The obtained results showed that on each of the three post-tests, the experimental groups outperformed the control group. After the treatment, they might improve their capacity for self-controlled learning, critical analysis, and problem-solving. The findings demonstrated that language learning for EFL learners can be enhanced by teaching through peer and self-assessment.

Our findings are consistent with Ashraf and Mahdinezhad's (2015) findings, which supported the merits of self- and peer assessment in fostering autonomy in language use with an emphasis on speaking ability. Additionally, Ritonga et al. (2022) confirmed the effects of peer assessment on enhancing Iranian EFL learners' reading comprehension, reading motivation, and vocabulary development, which supports the results of the study.

The results are further supported by Khairani (2019), who demonstrated the efficiency of task-based learning in improving reading skills through peer assessment. Additionally, Seifert and Feliks (2019) validated the benefits of self- and peer evaluation on the

improvement of students' assessment abilities and cognitive processes. Additionally, our findings are consistent with those of Tunagür (2021), who suggested a positive role of peer assessment in the writing motivation and anxiety of sixth-graders. Our results also align with those of Imani (2022), who showed the beneficial impacts of peer and self-evaluation on the speaking ability of EFL learners.

Moreover, Salem Almahasneh and Abdul-Hamid (2019), who examined the effects of peer assessment training on the writing skill of Arab EFL high school students, advocate the outcomes of our research. Their results indicated that the students who took part in peer assessment training outstripped on the writing post-test. Also, our results are in line with Patri (2002) whose results verified the helpful influences of peer- and self-assessments on college students of Chinese oral presentation skills. Besides, the findings gained in this research are supported by Abolfazli Khonbi and Sadeghi (2012) who showed that using self- and peer assessments generated positive effects on Iranian EFL learners' course achievement.

Similarly, our findings in this survey are reinforced by Li et al. (2021) who figured out the efficiency of peer assessment on EFL learners' non-cognitive skills. Furthermore, the existing investigation is verified by Lee (2015) who concluded that peer assessment can lead to more language improvement since learners are probably more interested in taking part in the assessment and learn more simply from their friends since they understand peer feedback better than teachers' feedback. The findings are in alignment with those of previous researchers, such as Kustati and Yuhardi (2014) and Topping (2018) who stated that peer assessment makes more progress in writing and enhances guided learning and students' problem-solving skills.

The findings of this study are congruent with the one conducted by Fathi and Khodabakhsh (2020), who found both self-assessment and peer assessment activities helped reduce the writing anxiety of the participants. Similarly, Fathi et al. (2019) analyzed the effect of practicing self-assessment and peer assessment activities on the L2 writing self-regulation of Iranian EFL learners. They argued that self-assessment and peer assessment activities made learners concentrate more on the demands of written tasks and found out how to develop their writing competencies and all of their linguistic resources to take more charge of their writing performance.

Early indications suggest that students who saw the benefits of peer evaluation enjoyed the sessions and had a deeper knowledge of the assessment, which supports the outcomes of the current study. Learners often benefit more from self- and peer assessment settings than from ones involving tutor-marked work. They learn knowledge through completing exams and frequently obtaining both oral and written feedback. To ensure that all students are treated fairly and equally, the tutor should monitor feedback and, if necessary, elaborate on it. The tutor can design the assessment criteria, but it adds more value when the students actively participate in the process. Students are engaged in the group and can demonstrate leadership abilities through peer assessment.

What is published in the literature can offer additional support for the conclusions of the present investigation. The presence of a competitive environment among the students during peer assessment and their desire to accurately assess their peers' performance may have prompted them to conduct a more thorough investigation and to be strict when creating measurement criteria and item construction, two factors that have

an impact on the effectiveness of assessment practice. Peer assessment reduces issues because it pairs people up and encourages them to concentrate on giving better performances at the next meeting. Peer evaluation can lessen the shortcomings that cause pupils to become anxious when they become stuck in their thinking.

Due to the effectiveness of both self- and peer assessment in fostering self-regulated learning, critical thinking, and problem-solving abilities in language use, particularly in speaking classes, the results of the present study have some pedagogical implications. Teachers may acquire the idea to instruct their pupils in a method that allows them to evaluate their own language learning and acquisition. One of the most important abilities they are supposed to acquire from universities and institutes is learning autonomy. Since the prerequisite for every sort of assessment is the assessor's profound understanding of the issue under study, students might profit from evaluating both themselves and their peers.

According to Tavakoli (2010), self-assessment would improve the student-teacher connection by empowering students to take ownership of their learning and progress, which would encourage them to participate more actively in their evaluation for the next learning goals. Because it motivates language learners to evaluate their learning progress and therefore aids them in maintaining their focus on their learning, it has been stated that self-assessment acts as an effective language learning approach to enhance autonomous language learning (Chen, 2005).

The cooperation fostered by peer evaluation among students could be one factor for the peer assessment group's superior performance as compared to the control group. Peer assessment can promote cooperative learning among students, as they are eager to assist and evaluate their colleagues, as well as take responsibility for their language learning achievement. This can lead to stronger social skills, more accurate evaluations, and better evaluations. Our findings are supported by the social interdependence theory, which states that students help one another learn better because they care about the group and its members and want to achieve the same objective (Slavin, 2011). According to the social interdependence theory, teamwork among students can help them achieve their common goals. Our findings also lend support to Vygotsky's social constructivism theory, which holds that cooperative learning activities benefit students by allowing them to work in each other's Zones of Proximal Development (ZPD) and observe how others behave, which is more beneficial than having them work alone (Webb, 2008).

Conclusion and implications

With the aid of this study, students can become more self- and peer-directed, and teachers can act as collaborators and facilitators. By observing their classmates, students can have a better grasp of how their peers learn. They are independent learners. They are more engaged in assessment and assume greater accountability for their learning. The findings of this study can provide insight to syllabus designers for creating a suitable syllabus. When creating a syllabus, syllabus designers should respect and take into account the learners' right to make their judgments and offer suggestions. The results of the current study encourage meaningful learning and lessen the issues associated with rote learning. Peer assessment was also demonstrated to be more beneficial for instructing and deep learning. Given that both evaluation methods were proven to be equally

helpful for EFL learners acquiring the language, syllabus designers are advised to include a variety of assessments kinds in their curricula.

Numerous things might be helpful in language learning. One of them could be evaluated. Students actively participate in the evaluation procedure and contribute to building language users' autonomy through self- and peer assessment. By supporting students in honestly appraising their own and their peers' accomplishments, self- and peer assessment promotes lifelong learning. Peer assessment allows students to participate in an important component of higher education while also providing a critical appraisal of the work of others, making them feel like they are a member of an academic community.

Providing additional feedback from peers while enabling teachers to evaluate individual pupils less, but better, is a crucial function of self- and peer evaluation. This creates a shift in the evaluation of student work from quantity to quality and also the higher-order thinking abilities. The study demonstrates the necessity of using self- and peer assessment in problem-based learning materials. The results of this study demonstrate the clear benefits of self- versus peer assessment in helping EFL learners develop their self-regulated learning, critical thinking, and problem-solving abilities.

The first step in training students to become successful responders is to make them aware of the value of providing effective feedback. It is considerably more important to teach kids how to reply to their classmates' writing. Here are some things teachers may do to help students with their assessments:

- Tell pupils to concentrate on specific areas of their writing. Students should concentrate on the content of the essay (how the concepts and thoughts are organized) rather than the grammatical or formal faults in the first draught. Students should concentrate more on formal faults (e.g., grammar, expression, sentence structure) in future writing.
- Teachers should give pupils directional surveys to evaluate their peers' work. This is considered the most important duty, so students can rely on it to provide acceptable comments to their friends based on the teacher's orientation. Another thing to keep in mind is that each form of writing necessitates the use of a unique response-oriented questionnaire. Teachers should avoid utilizing generic questionnaires to ensure that students' feedback is general and does not delve into the precise substance of each paper. Teachers having feedback tasks for students to work independently, in pairs, or in groups to decide whether a piece of writing is good or not is one of the most successful feedback activities. This way, students can know what aspect to focus on when giving feedback and what issues they need to keep in mind when writing a paragraph.

To elaborate on the pedagogical implications of this study, it can be argued that because the current study's findings demonstrated that using peer and self-assessments can have positive effects on learners' self-regulated learning, critical thinking, and problem-solving skills, some language teaching and learning activities can be designed. It is strongly advised that L2 policymakers, curriculum creators, syllabus designers, teacher educators, and test developers take an alternate assessment on the board more seriously. The incorporation of peer and self-assessments in language

instruction can help EFL teachers assign a more responsible and autonomous role to their students by delivering more motivating and self-regulated learning in induced and sustained learning.

Based on the findings, we can conclude that this study can provide some benefits to students during the learning process. Students, for example, might boost their incentive to attain learning objectives through self-assessment. They can also improve their communication with the teachers. Furthermore, self-assessment might encourage students to learn independently in order to meet learning objectives and enhance their talents for future performances. Furthermore, the findings of this study may help students identify their strengths and weaknesses in English language acquisition. This study helps students become peer and self-directed, and enables teachers to be facilitators and collaborators. Students are able to witness their peers and arrive at a better understanding of how their peers learn. They are autonomous learners. They take more responsibility in their own learning, and have more involvement in assessment.

To take advantage of self-evaluation and peer assessment, teachers may want to devote sufficient theoretical and practical attention to these alternative assessment methodologies. By doing so, it is ensured that Iranian EFL teachers gain a thorough awareness of the benefits and drawbacks of self-assessment and peer assessment and, as a result, employ them in their actual lessons. Language teachers are strongly encouraged to incorporate more instructional practices such as self- and peer assessment into their instruction, which may ensure students' learning and increase their motivation and independence, both of which are key aspects in the learning process.

Research limitations and suggestions

Although we tried to write a perfect research, we could not escape from some limitations. We could include only 75 participants in our research. Only quantitative data were collected to answer the research questions. We could train the participants in only 13 conversations in 13 sessions. Because of gender segregation, we could work on male students.

Finally, several recommendations for additional research are made. Regarding the treatment phase, it is advised to repeat this investigation over a longer length of time. The same study can be done on a larger sample of EFL students to provide the researcher(s) with more precise and broadly applicable results. It is advised that other studies make use of different rating scales and self- and peer assessment approaches.

Abbreviations

EFL	English as Foreign Language
ANOVA	Analysis of variance
OQPT	Oxford Quick Placement Test
CEFR	Common European Framework of Reference for Languages
SILL	Strategy Inventory for Language Learning
CT	Critical thinking
PSS	Problem-solving skills

Acknowledgements

Not applicable.

Authors' contributions

All authors had equal substantial contributions to conception and design, acquisition of data, analysis and interpretation of data, and writing the manuscript. All authors read and approved the final manuscript.

Authors' information

Dr. Seyyed Mohammad Ali Soozandehfar is an Assistant Professor of TEFL in the Department of Teaching English as a Foreign Language (TEFL) at University of Hormozgan, Iran. He received his BA in English Language and Literature, MA and PhD both in TEFL, in 2008, 2011, and 2015, respectively, all from Shiraz University.

Dr. Tribhuvan Kumar is an assistant professor at the Department of English Language and Literature, College of Science and Humanities at Sulail, Prince Sattam Bin Abdulaziz University, Al Kharj, Saudi Arabia.

Dr. Arash Hashemifardnia got his PhD in 2019 from IAU of Shahrekord. He is now a teacher teaches at high schools in Ahvaz, Iran.

Mr. Rasoul Mombeini got his MA from the English Department, Ahvaz Branch, Islamic Azad University, Ahvaz, Iran

Funding

No funding received in this study.

Availability of data and materials

The authors state that all the data supporting the findings of this study are available within the article.

Declarations**Competing interests**

The authors declare that they have no competing interests.

Received: 29 May 2023 Accepted: 13 July 2023

Published online: 24 July 2023

References

- Abolfazli Khoonbi, Z., & Sadegh, K. (2012). The effect of assessment type (self vs. peer vs. teacher) on Iranian university EFL students' course achievement. *Language Testing in Asia*, 2(4), 47–74.
- Adachi, C., Tai, J. H. M., & Dawson, P. (2018). Academics' perceptions of the benefits and challenges of self and peer assessment in Higher Education. *Assessment and Evaluation in Higher Education*, 43(2), 294–306.
- Alias, M., Masek, A., & Salleh, H. H. M. (2015). Self, peer, and teacher assessments in problem-based learning: are they in agreements? *Procedia-Social and Behavioral Sciences*, 204, 309–317.
- Allen, D. (2004). *Oxford placement test 1*. OUP.
- Argaw, A. S., Haile, B. B., Ayalew, B. T., & Kuma, S. G. (2017). The effect of problem based learning instruction on students' motivation and problem solving skills of physics. *Eurasia Journal of Mathematics, Science and Technology Education*, 13(3), 857–871.
- Ashraf, H., & Mahdinezhad, M. (2015). The role of peer-assessment versus self-assessment in promoting autonomy in language use: a case of EFL learners. *Iranian Journal of Language Testing*, 5(2), 110–120.
- Azizi, Z., & Namaziandost, E. (2023). Implementing peer-dynamic assessment to cultivate Iranian EFL learners' inter-language pragmatic competence: a mixed-methods approach. *International Journal of Language Testing*, 13(1), 18–43. <https://doi.org/10.22034/ijlt.2022.345372.1171>
- Babaii, E., Taghaddomi, S., & Pashmforoosh, R. (2016). Speaking self-assessment: Mismatches between learners' and teachers' criteria. *Language Testing*, 33(3), 411–437.
- Birjandi, P., & Bolghari, M. S. (2015). The relationship between the accuracy of self-and peer-assessment of Iranian intermediate EFL learners and their learning styles. *Theory and Practice in Language Studies*, 5(5), 1–18.
- Black, P., & William, D. (1998). *Inside the black box: raising standards through classroom assessment*. Department of Education and Professional Studies, Kings College.
- Bobkina, J., & Stefanova, S. (2016). Literature and critical literacy pedagogy in the EFL classroom: Towards a model of teaching critical thinking skills. *Studies in Second Language Learning and Teaching*, 6(4), 677–696.
- Boekaerts, M., & Corno, L. (2005). Self-regulation in the classroom: A perspective on assessment and intervention. *Applied Psychology: An International Review*, 54(2), 199–231. <https://doi.org/10.1111/j.1464-0597.2005.00205.x>
- Boud, D. (2000). Sustainable assessment: rethinking assessment for the learning society. *Studies in Continuing Education*, 22, 151–16.
- Brown, H. D., & Abeywickrama, P. (2010). *Language assessment: principles and classroom practices (Vol. 10)*. Pearson Education: White Plains.
- Brown, J. D., & Hudson, T. (2012). The alternatives in language assessment. *TESOL Quarterly*, 32(4), 653–675.
- Butler, Y., & Lee, J. (2010). The effects of self-assessment among young learners of English. *Language Testing*, 27(1), 57–69.
- Chen, Y. (2005). *Peer and self-assessment in EFL classrooms. Selected Papers of the Fourteenth International Symposium on English Teaching* (pp. 320–329). Taipei: Crane Publishing.
- Chen, Y. (2008). Learning to self-assess oral performance in English: A longitudinal case study. *Language Teaching Research*, 12(1), 235–262.
- Cheng, W., & Warren, M. (2005). Peer assessment of language proficiency. *Language Testing*, 22(1), 93–121.
- Chua, B. L., Tan, O. S., & Liu, W. C. (2016). Journey into the problem-solving process: cognitive functions in a PBL environment. *Innovations in Education and Teaching International*. <https://doi.org/10.1080/14703297.2014.961502>
- Clark, I. (2012). Formative assessment: assessment is for self-regulated learning. *In Educational Psychology Review*. <https://doi.org/10.1007/s10648-011-9191-6>
- Davies, P. (2006). *Peer-assessment: judging the quality of students' work by comments rather than marks. Innovations in Education and Teaching International*, 43, 69–82.

- Docktor, J., Heller, K. (2009). Robust assessment instrument for student problem solving. *Proceedings of the NARST 2009 Annual Meeting*. <https://doi.org/10.1063/1.3266696>
- Esfandiari, S., & Tavassoli, K. (2019). The comparative effect of self-assessment vs. peer-assessment on young EFL learners' performance on selective and productive reading tasks. *Iranian Journal of Applied Linguistics (IJAL)*, 22(2), 1–35.
- Falchikov, N. (2005). *Improving assessment through student involvement. Practical solutions for aiding learning in higher and further education*. Routledge Falmer.
- Fathi, J., & Khodabakhsh, M. R. (2020). Self-Assessment and peer-assessment in writing course of Iranian EFL students: An investigation of writing anxiety. *International Journal of English Language and Translation Studies*, 8(1), 88–96.
- Fathi, J., Mohebiniya, S., & Nourzadeh, S. (2019). Enhancing second language writing self-regulation through self-assessment and peer-assessment: a case of Iranian EFL learners. *International Journal of Applied Linguistics and English Literature*, 8(3), 110–117.
- Fitriani, A., Zubaidah, S., Susilo, H., & Al Muhdhar, M. H. I. (2020). The effects of integrated problem-based learning, predict, observe, explain on problem-solving skills and self-efficacy. *Eurasian Journal of Educational Research*, 85, 45–64.
- Fu, Q. K., Lin, C. J., & Hwang, G. J. (2019). Research trends and applications of technology-supported peer assessment: A review of selected journal publications from 2007 to 2016. *Journal of Computers in Education*, 6(2), 191–213.
- Fulcher, G., & Davidson, F. (2007). *Language testing and assessment: An advanced resource book*. London: Routledge.
- Ghahderijani, B. H., Namaziandost, E., Tavakoli, M., Kumar, T., Magizov, R. (2021). The comparative effect of group dynamic assessment (GDA) and computerized dynamic assessment (C-DA) on Iranian upper-intermediate EFL learners' speaking complexity, accuracy, and fluency (CAF). *Language Testing in Asia*, 11(1). <https://doi.org/10.1186/s40468-021-00144-3>.
- Hadwin, A. F., Jarvela, S., & Miller, M. (2015). Self-regulated, co-regulated, and socially shared regulation of learning. In *Handbook of Self-Regulation of Learning and Performance*. <https://doi.org/10.4324/9780203839010.ch5>
- Han, H.X.D., Toh, T.L. (2019). Use of animation to facilitate students in acquiring problem-solving: From Theory to Practice. *The Mathematics Enthusiast*, 16(1), 377–388. <https://doi.org/10.54870/1551-3440.1464>.
- Harris, L. R., & Brown, G. T. (2013). Opportunities and obstacles to consider when using peer-and self-assessment to improve student learning: case studies into teachers' implementation. *Teaching and Teacher Education*, 36, 101–111.
- Hedge, T. (2000). *Teaching and learning in language classroom*. Oxford University Press.
- Hong Canh, T. (2022). The effects of peer assessment on EFL students' writing performance at a lower secondary school in Tien Giang Province – Vietnam. *OSR Journal Of Humanities And Social Science (IOSR-JH)*, 27(8), 9–24.
- Hu, Y., Wu, B., & Gu, X. (2017). Learning analysis of K-12 students' online problem solving: A three stage assessment approach. *Interactive Learning Environments*, 4820(March), 262–279. <https://doi.org/10.1080/10494820.2016.1276080>
- Hung, Y. J. (2018). Group peer assessment of oral English performance in a Taiwanese elementary school. *Studies in Educational Evaluation*, 59, 19–28.
- Hung, Y. (2019). Bridging assessment and achievement: repeated practice of self-assessment in college English classes in Taiwan. *Assessment and Evaluation in Higher Education*, 44(8), 1191–1208.
- Hung, Y. J., Samuelson, B. L., & Chen, S. C. (2016). Relationships between peer- and self-assessment and teacher assessment of young EFL learners' oral presentations. In M. Nikolov (Ed.), *Assessing young learners of English: Global and local perspectives* (pp. 317–338). Springer.
- Imani, S. (2022). The comparative effect of self-assessment and peer assessment on reflective and impulsive EFL learners' speaking skill. *Journal of Foreign Language Teaching and Translation Studies*, 6(4), 99–120.
- Khairani, A. (2019). *The implementation of peer assessment in reading skill through task based learning at grade tenth in SMA Negeri 1 Labuhan Deli*. A thesis: English Department. Faculty of Languages and Arts. State University of Medan
- Kuiper-Anne, R., & Pesut, J. D. (2016). Promoting cognitive and metacognitive reflective reasoning skills in nursing practice: self-regulated learning theory. *Journal of Advance Nursing*, 3, 19–34.
- Kustati, M., & Yuhardi. (2014). The effect of the peer-review technique on students' writing ability. *Studies in English Language and Education*, 1(2), 80–90.
- Laurens, T., Batlolona, F. A., Batlolona, J. R., & Leasa, M. (2018). How does realistic mathematics education (RME) improve students' mathematics cognitive achievement? *Eurasia Journal of Mathematics, Science and Technology Education*, 14(2), 569–578.
- Lee, M. K. (2015). Peer feedback in second language writing: Investigating junior secondary students' perspectives on inter-feedback and intra-feedback. *System*, 55, 1–10.
- Li, H., Bialo, J., Xiong, Y., Vincent Hunter, C., & Guo, X. (2021). The effect of peer assessment on non-cognitive outcomes: a meta-analysis. *Applied Measurement in Education*. <https://doi.org/10.1080/08957347.2021.1933980>
- Ma, W., & Winke, P. (2019). Self-assessment: how reliable is it in assessing oral proficiency over time? *Foreign Language Annals*, 52(1), 66–86.
- Marzuki, A., Alek, A., Farkhan, M., Deni, R., & Raharjo, A. (2020). Self-Assessment in exploring EFL students' speaking skill. *Al-Ta'lim Journal*, 27(2), 206–212.
- Matsuno, S. (2009). Self-, peer-, and teacher-assessments in Japanese university EFL writing classrooms. *Language Testing*, 26, 75–100.
- McKay, P. (2006). *Assessing young language learners*. Cambridge University Press.
- Meletiadiou, E. (2021). *Exploring the impact of peer assessment on EFL students' writing performance*. London South Bank University.
- Miller, P. J. (2003). The effect of scoring criteria specificity on peer and self-assessment. *Assessment and Evaluation in Higher Education*, 28(4), 383–394.
- Musfirah, Y. (2019). The use of peer assessment in speaking skill. *English Education Journal*, 10(1), 67–79.
- Namaziandost, E., Heydarnejad, T., Azizi, Z. (2023). To be a language learner or not to be? The interplay among academic resilience, critical thinking, academic emotion regulation, academic self-esteem, and academic demotivation. *Current Psychology*. <https://link.springer.com/article/https://doi.org/10.1007/s12144-023-04676-0>.
- Otite, O., & Ogionwo, W. (2006). *An introduction to sociological Studies*. Ibadan: Heinemann Educational Books (Nigeria) Plc.
- Oxford, R. L. (1990). *Language learning strategies: what every teacher should know*. Boston: Heinle & Heinle.

- Patri, M. (2002). The influence of peer feedback on self-and peer assessment of oral skills. *Language Testing*, 19(2), 109–131.
- Paul, R., & Elder, L. (2008). *The miniature guide to Critical Thinking concepts and tools: Foundation for Critical Thinking Press*.
- Pedrotti, M., Nistor, N. (2019). How Students Fail to Self-regulate Their Online Learning Experience (M. Scheffel, J. Broisin, V. Pammer-Schindler, A. Ioannou, & J. Schneider, Eds.). Springer Link; Springer International Publishing. https://doi.org/10.1007/978-3-030-29736-7_28.
- Rezai, A., Namaziandost, E., & Amraei, A. (2023). Exploring the effects of dynamic assessment on improving Iranian quran learners' recitation performance. *Critical Literary Studies*, 5(1), 159–176. <https://doi.org/10.34785/J014.2023.010>
- Rezai, A., Namaziandost, E., Miri, M., & Kumar, T. (2022). Demographic biases and assessment fairness in classroom: insights from Iranian university teachers. *Language Testing in Asia*, 12(1). <https://doi.org/10.1186/s40468-022-00157-6>
- Richards, J. C., & Schmidt, R. (2002). *Dictionary of language teaching and applied linguistics*. Pearson Education Limited.
- Ritonga, M., Tazik, K., Omar, A., & Dehkordi, E. (2022). Assessment and language improvement: The effect of peer assessment (PA) on reading comprehension, reading motivation, and vocabulary learning among EFL learners. *Language Testing in Asia*, 5, 12–36.
- Roberts, T. (2006). *Self-, peer-, and group assessment in E-learning*. Information science publishing.
- Saito, H. (2008). EFL classroom peer-assessment: training effects on rating and commenting. *Language Testing*, 25, 553–581.
- Saito, H., & Fujita, T. (2009). Peer-assessing peers' contribution to EFL group presentations. *RELC Journal*, 40(2), 149–171.
- Salem Almahasneh, A., & Abdul-Hamid, S. (2019). The effect of using peer assessment training on writing performance among Arab EFL high school students in Malaysia. *Arab World English Journal (AWEJ)*, 10(1), 105–115.
- Seker, M. (2015). The use of self-regulation strategies by foreign language learners and its role in language achievement. *Language Teaching Research*, 20, 600–618. <https://doi.org/10.1177/1362168815578550>.
- Slavin, R. E. (1997). *Educational psychology: theory and practice*. MA: Allyn & Bacon: Needham Heights.
- Slavin, R. E. (2011). *Student Team Learning: A Practical Guide to Cooperative Learning* (3rd ed.). Washington DC, National Education Association.
- Schunk, D.H., Zimmerman, B.J. (2012). Motivation and self-regulated learning: theory, research, and applications. In *Motivation and Self-Regulated Learning: Theory, Research, and Applications*. <https://doi.org/10.4324/9780203831076>
- Seifert, T., & Feliks, O. (2019). Online self-assessment and peer-assessment as a tool to enhance student-teacher assessment skills. *Assessment and Evaluation in Higher Education*, 44(2), 169–185.
- Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 77–84.
- Sukariasih, L., Tahang, L., Nursalam, L.O., & Fayanto, S. (2020). Description of physics problem-solving in the topic of static fluid: Case study of physics education in Halu Oleo University. *Universal Journal of Educational Research*. <https://doi.org/10.13189/ujer.2020.081025>
- Tavakoli, M. (2010). Investigating the relationship between self-assessment and teacher-assessment in academic contexts: A Case of Iranian university students. *The Asian EFL Journal Quarterly*, 12(1), 234–260.
- Thomas, F., & Nelson, L. (2010). *About the California critical thinking disposition inventory*. Indiana University.
- Tigchelaar, M. (2019). Exploring the relationship between self-assessments and OPIc ratings of oral proficiency in French. *Dalam Foreign language proficiency in higher education* (hlm. 153–173). Springer.
- Topping, K. (2009). Peer-assessment. *Theory into Practice*, 48, 20–27. <https://doi.org/10.1080/00405840802577569>.
- Topping, K. J. (2018). *Using peer assessment to inspire reflection and learning*. London, Routledge.
- Tunagür, M. (2021). The effect of peer assessment application on writing anxiety and writing motivation of 6th grade students. *International Journal of Education*, 10(1), 96–105.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, Harvard University Press.
- Watson, G. B., & Glaser, E. M. (1991). *Watson-Glaser Critical Thinking Appraisal Manual: British Manual: Forms A, B and C*. The Psychological Corporation: Harcourt Brace.
- Watson G. B., & Glaser, E. M. (2002). *Watson-Glaser Critical Thinking Appraisal Manual*. The Psychological Corporation: Harcourt Brace.
- Webb, N. M. (2008). Learning in small groups. In T. L. Good (Ed.), *21st Century education: a reference handbook* (pp. 203–211). Sage.
- Weisi, H., & Karimi, M. (2013). The effect of self-assessment among Iranian EFL learners. *Procedia - Social and Behavioral Sciences*, 70(25), 731–737.
- Wong, J., Baars, M., de Koning, B. B., & Paas, F. (2021). Examining the use of prompts to facilitate self-regulated learning in massive open online courses. *Computers in Human Behavior*. <https://doi.org/10.1016/j.chb.2020.106596>
- Yan, Z., Wang, X., Boud, D., & Lao, H. (2023) The effect of self-assessment on academic performance and the role of explicitness: a meta-analysis. *Assessment & Evaluation in Higher Education*, 48(1), 1–15. <https://doi.org/10.1080/02602938.2021.2012644>
- Zhi-Feng, E., & Yi LEE, C. (2013). Using peer feedback to improve learning via online peer-assessment. *Turkish Online Journal of Educational Technology*, 12(1), 187–199.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.