Peer assessment in group-oriented classroom contexts: on the effectiveness of peer assessment coupled with scaffolding and group work on speaking skills and vocabulary learning

Mina Homayouni[*]

*Correspondence:
mina.homayouni89@yahoo.com

Department of English
Language and Literature, Faculty
of Humanities, Shahid Chamran
University of Ahvaz, Ahvaz, Iran

Abstract

Peer learning, also known as collaborative learning, is based on social constructivism and contends that learning takes place more actively when students interact socially with their peers. This study sought to examine the effects of scaffolded peer assessment in group-oriented classrooms on developing speaking skills and enlarging the vocabulary size of language learners. To accomplish this objective, through cluster sampling, the researchers selected 37 lower-intermediate and 5 intermediate learners of English. Then, 20 lower-intermediate subjects were randomly assigned as the experimental group along with the 5 more proficient learners. In groups of 5, the intermediate learner was assigned the role of the mediator and was responsible for giving feedback to their peers. The rest of the subjects were assigned as the control group and there was no mediator in their group. Four instructional sessions were allocated to scaffolded peer assessment of speaking and four sessions were devoted to scaffolded peer assessment of vocabulary learning. In this randomized pre-test–post-test-delayed post-test trial, an independent sample *t* test, and a one-way repeated measures ANOVA were carried out. The results of the statistical analyses demonstrated the impact of scaffolded peer assessment on developing both speaking skills and enlarging learners' vocabulary size with a large effect size. That is, by implementing scaffolded peer assessment in a group-oriented context both speaking skills and vocabulary knowledge can be developed. The pedagogical implication of this study is that language teacher can implement the notions of social constructivism and socio-cultural theory proposed by Vygotsky (Readings on the Development of Children 23:34-41, 1978) to expand and develop learners' zone of proximal development.

Keywords: Peer assessment, Scaffolding, Speaking skills, Vocabulary learning, Zone of proximal development

Introduction

Standardized tests, teacher evaluations, and examinations, among other traditional methods of assessment, have recently fallen out of favor since they do not place enough emphasis on students' responsibility (McNamara, 2001). Teachers are currently urged

to "collaborate with their students in terms of the employment of peer and self-assessments" in higher education (Pantiwati & Husamah, 2017, p. 187). Peer assessment has drawn a lot of attention recently because it is pedagogically significant and places a greater emphasis on the independence and autonomy of the learner (Patri, 2002). It is commonly known that peer evaluation "contributes to the growth of student learning and encourages ownership of assessment processes" (Bryant & Carless, 2010, p. 3). This is true even if students have described peer assessment as challenging (Kearney, 2013).

Feedback is frequently seen as a key element of educational practice and essential to students' growth and learning (Fyfe & Rittle-Johnson, 2016). One strategy for providing learners with feedback properly has been recognized as peer assessment (Van Zundert et al., 2010). Peer assessment, peer feedback, peer evaluation, and peer grading are some of the terms used in the literature to describe the use of students to generate comments regarding their peers' performances (Double et al., 2020). This type of feedback can be delivered in a variety of forms including verbal feedback, written comments, and/or grades (Topping, 1998). Critically speaking, students' learning may gain more by taking on both the position of assessor and being assessed than it may from being assessed only (Reinholz, 2016).

Peer evaluation is a communication technique that allows students to have conversations about their own performance and academic requirements (Liu & Carless, 2006, p. 280). Peer assessment can be used in second language curricula and courses as a form of formative assessment and collaborative learning. Peer assessment can give authors the chance to discuss their texts and learn other people's interpretations of them (Hyland, 2000), improve students' production skills by incorporating peer assessment in revisions (Zhao, 2010), spark students' interest in production (Shih, 2010), scaffold students' production process, and improve critical thinking (Joordens et al., 2009). Peer assessment can also support learner autonomy, which is more significant (Yang et al., 2006).

The belief that interactions with others might have a significant impact on a learner's growth naturally sparks curiosity about what those others might accomplish. Centered on supporting young children's language acquisition through scaffolding, it is a "process that enables a kid or novice to solve problems, carry out a task, or attain a goal which would be beyond his unsupported efforts," according to Wood et al. (1976, p. 90). Scaffolding has been viewed as a social interaction from the beginning, thus focusing on the participants in that interaction was a logical next step, especially for sociocultural theorists with roots in the Vygotskian school of thought (Wright, 2018).

In circumstances involving second or foreign languages, interaction can be seen as a crucial component of language acquisition. According to Ellis (2015), interaction is the discourse that the sender and the receiver both form and which results in input. As a result, the interactionist paradigm is predicated on the idea that interaction between the learners' mental capabilities and the linguistic context leads to language learning (Long, 1996). Long claims that the three aspects of language interaction are input, production, and feedback. Input is the language that is given to learners by native speakers, teachers, or peers; production is the language that learners speak for themselves; and feedback is the response that interlocutors give to the learners' production (Khezrlou et al., 2017).

According to the sociocultural theory put forth by Vygotsky in 1978, social interaction is the optimal environment for learning. According to him, this form of learning

promotes growth. Based on this viewpoint, social engagement offers the best kind of feedback for students' mistakes. The learner receives the corrective feedback based on his or her zone of proximal development (ZPD) through this type of interaction rather than receiving explicit or implicit corrective feedback. Instead, the learner moves from the most implicit to the most explicit corrective feedback through a step-by-step and contingent feedback negotiation (Nassaji & Swain, 2000). ZPD is the distance between what the learner can do independently and what s/he will be able to do with the help of a more capable other (Fulcher & Davidson, 2007). In other words, the distance between the learner's actual and potential ability level is called ZPD. The learners are placed in a problem-solving process throughout this language negotiation process instead of being given the correct response, allowing them to transition from other-regulation to self-regulation (Aljaafreh & Lantolf, 1994). As learners go along this continuum, their self-esteem increases and language evaluation becomes ingrained in their language proficiency, resulting in cognitive development (Ahmadpour & Yousefi, 2016; Heshmat Ghahderijani et al., 2021).

In addition to what went above, speaking ability plays a crucial role in learning English. Speaking is one of the signs that someone has mastered a language because it allows one to express thoughts and ideas through spoken language (Fauzan, 2014; Jiang et al., 2022). According to Ur (1996:120), speaking appears to be the most crucial skill since those who are fluent in a language are called "speakers," and the majority of people who are studying a foreign language are primarily interested in speaking the language. Nunan (1999) contends that speaking the target language is the most common way to define the capacity to function in a foreign language. Given the significance of speaking proficiency, several English teachers concentrate on the best ways to instruct speaking in their classes. Additionally, According to Elleman and Oslund (2019), vocabulary growth is a trustworthy sign of linguistic progress. Students who lack sufficient vocabulary knowledge endure difficulties and setbacks throughout their language-learning careers, which demotivates them and causes them to fail (Khany & Khosravian, 2014; Namaziandost et al., 2022). It is for these reasons that speaking skills and vocabulary learning need to be developed, especially because students can improve their vocabulary knowledge by being helped and scaffolded (Vadivel et al., 2021).

Numerous kinds of research that put Vygotskian ideas through empirical analysis have demonstrated the effectiveness of this strategy in the growth of linguistic competence. However, there are not many studies on peer assessment, and certainly not on peer assessment of oral skills, despite the fact that it has been highlighted as a key learning lever (Gielen et al., 2010). Studies using verbal peer feedback on oral presentations are far less prevalent than studies that take the form of written feedback on writing (e.g., Cartney, 2010; Yang et al., 2006). Furthermore, there is not enough scaffolding research that is only concerned with situations in higher education (Doo et al., 2020). Thus, there is a need to look into how peer assessment and scaffolding in group-oriented classrooms affect language learners' oral performance and their vocabulary retention and recall. Accordingly, this study is an attempt to delve into Vygotskian thinking and check the extent to which applying his theories coupled with peer assessment and scaffolding can facilitate oral skill and/or vocabulary recall and retention.

It is widely accepted that vocabulary constitutes the building block of language. Additionally, nobody denies that the ultimate goal of a large number of language learners is to be able to speak the language fluently. Therefore, attempts need to be made to make learners both fluent in English and also increase the vocabulary at their disposal. One pedagogical technique that has attracted researchers' attention is group-oriented learning based on the principles of cooperative learning which draw on Vygotskian thinking. It is hoped that the results of this study will add to the literature and make it crystal clear whether by implementing Vygotskian thinking in a practical situation learners' speaking skills and vocabulary learning can be affected. Furthermore, it is also hoped that policy-makers, materials developers, etc. will also benefit from the findings of this study.

Literature review

Theoretical underpinnings

Peer assessment

Giving students evaluation experiences similar to those of their teachers, such as peer feedback, which is widely acknowledged as a fruitful context for improving engagement and developing evaluative competencies, is one strategy for developing feedback-enabling processes (Cartney, 2010). Peer feedback is said to be a formative assessment activity that concentrates on qualitative comments, although it is sometimes used interchangeably with peer assessment or peer review (Liu & Carless, 2006).

Additionally, assessment can help instructors by assisting them in identifying the strengths and weaknesses of their students (Baniabdelrahman, 2010). Additionally, it is said that evaluation is regarded as an honest, ethical, and successful tool for evaluating various student characteristics (Mousavi, 2012). Peer assessment is a typical evaluation kind that fosters learners' critical thinking and creativity abilities. While evaluating each other's writing abilities, students also grow their own learning (Chong et al., 2012). In general, PA can be defined as the process by which students' work is evaluated by their peers based on predetermined standards (Tunagür, 2021). A series of acts known as PA can be used by students to evaluate and compare the learning efforts and successes of their peers (Liu & Brantmeier, 2019). Researchers have noted that PAs have certain positive characteristics, such as motivating students, enhancing involvement, and giving students ownership of their learning (Sluijsmans, 1999; Xu et al., 2022).

"A communication technique through which learners engage in discussions about performance and standards" is referred to as peer assessment (Liu & Carless, 2006, p. 280). Peer evaluation is a process whereby students evaluate and specify the standard, worth, or caliber of output or performance of other students who are on an equal footing (Topping, 2009). Writing, oral presentations, portfolios, test performance, and other skillful activities are examples of products that can be peer-evaluated. Although it is more constrained, peer assessment can be seen as a form of collaborative learning (Falchikov, 2007; Syairofi et al., 2022). Simply put, it implies that students evaluate each other's work using pertinent standards and provide criticism for both the recipient's benefit and their own personal growth.

The inclusion of a peer assessment component in a course assessment diet has the potential to improve student engagement, accountability, and excellence while also establishing clearer course frameworks, emphasizing skills and learning, and

increasing feedback (Rezai et al., 2022; Weaver & Cottrell, 1988). By involving students in evaluating the work of their peers, peer assessment plays a crucial role in formative assessment. With appropriate implementation, it can also be used as a component in summative assessment. It can also be viewed as a method of learning in and of itself, as well as a mechanism for evaluating the outcomes of student learning. Less is known about students' opinions on assessing and being assessed by peers, even though these concepts about the beneficial role peer assessment can play in classroom assessment are well-known.

Peer assessment, the process in which learners can comment on the performance of their classmate, is a procedure that has been acknowledged as having significant potential benefits for learning gain and is increasingly being utilized in higher education to get students more involved in the assessment process (Race et al., 2005). Accordingly, peer feedback can play a decisive role in both language learning and formative assessment.

In short, peer assessments supporting student-centered assessments have attracted a lot of attention (Hoffman, 2019). Based on this style of evaluation, the students rank the performances of their classmates on a quantitative level, for example, by giving peer grades or scores, and/or a qualitative level, for example, by giving a peer oral or written feedback (Topping, 2017). By letting students evaluate their peers' performances during assessments, peer assessments promote significant educational ideals (Bryan & Clegg, 2019). Learner-centered evaluation and learning can be facilitated through peer assessment (Panadero et al., 2016; Tomas & Palmer, 2018).

Cooperative learning

Peer learning, also known as collaborative learning, is based on social constructivism and contends that learning takes place more actively outside of the classroom when students interact socially with their peers (Roschelle & Teasley, 1995). Learners positively rely on one another and help each other's mental models by sharing personal experiences, insights, and reflections (Johnson & Johnson, 1987; Slavin, 1982).

In a setting of cooperative learning, group members try to contribute individually to advance learning and reach a group goal (Johnson et al., 2014). This method aids in the cooperative knowledge-building of students (Naserpour & Zarei, 2021). In reality, there is a strong interdependence among group members while each person accepts responsibility for his or her own learning (Bolukbas et al., 2011).

According to Wajnryb (1990), when students collaborate to develop language, they become more aware of gaps in their linguistic knowledge and consider form-meaning linkages when they get peer feedback. In actuality, cooperative learning projects provide for the exposure of all members to a variety of learning benefits, methods, and strategies, allowing for the exhibition of each member's unique learning preferences or strengths (Naserpour & Zarei, 2021). Consequently, cooperative activities can help learners improve their language skills as well as their interpersonal abilities, such as teamwork and problem-solving (Olaya & González-González, 2020). The rich literature on collaborative learning provides strong evidence in terms of substantial gains in problem-solving and critical-thinking skills (Kyndt et al., 2013).

Scaffolding

The support given during the educational process to suit students' requirements when they are exposed to novel concepts and skills is referred to as scaffolding in Sawyer's (2006) work. Higher and more in-depth levels of learning may result from this (Naserpour & Zarei, 2021).

ZPD, a crucial concept in Socio-Cultural Theory, and scaffolding have a strong relationship. ZPD, as defined by Vygotsky (1978), is the discrepancy between a child's actual developmental level and their potential level of development, which is determined by how well they can solve problems when receiving advice from adults or more experienced peers (Verenikina, 2008). According to Vygotsky's social constructionism theory, scaffolding refers to a novice's temporary aid from an expert designed to increase their independence. As learners gain the capacity to exhibit mastery and carry out tasks on their own and develop their own abilities and talents, this assistance is gradually reduced or eliminated (Diaz-Rico & Weed, 2002).

To satisfy the needs of students in English language classrooms, scholars have suggested a variety of categories of scaffolding tactics. The most popular and useful scaffolding strategies that can be used in language classrooms to teach single and multi-words are categorized into three primary groups, according to Echevarria et al. (2004): verbal, procedural, and instructional, each of which includes several tactics.

Empirical background

Studies on the efficacy of peer assessment and scaffolding

One of the latest studies that have delved into the efficacy of peer assessment on developing speaking skills is that of Imani (2021). The purpose of her study was to compare how self- and peer-assessment affected reflective and impulsive EFL learners' speaking abilities. To do this, 51 intermediate EFL students (24 impulsive and 27 reflective) were chosen based on their performance on the PET and the Impulsiveness Sub-Scale of Eysenck's Impulsivity Inventory. They were split into two experimental groups at random—self-assessment and peer assessment—each of which included both impulsive and reflective students. The Speaking Self-Assessment Sheet created by Babaii et al. (2016) served as the basis for the evaluation in the self-assessment group (2015). In the peer assessment group, the evaluation was based on the Yamashiro and Johnson Peer Assessment Rating Sheet (Yamashiro & Johnson, 1997). The speaking portion of another PET was administered to the subjects as a post-test following the conclusion of the treatment. The two-way ANOVA results showed that (a) there was no significant interaction between assessment type and cognitive type, (b) self- and peer-assessment had the same effect on impulsive and reflective learners' speaking skills (i.e., they both similarly helped learners develop their speaking skill), and (c) regardless of assessment type, learners with different cognitive types performed differently, with reflective learners outperforming their impulsive counterparts.

Salem Almahasneh and Abdul- Hamid (2019) examined the impact of peer evaluation training on the writing performance of Arab EFL high school students. To carry out this research, 120 students between the ages of 15 and 16 participated in this study. The pupils were from two Arab high schools in Malaysia. Data collection and evaluation of the students' writing performance on the pre-test and post-test were done using an

analytical marking scale. The results of this study showed that the writing skills of the experimental and control groups were significantly different. The findings showed that students who participated in peer assessment training produced writing drafts that were superior to those of students who only received conventional essay writing instruction and no peer assessment.

Tunagür (2021) aimed to ascertain whether the application of peer assessment affected sixth-grade pupils' writing motivation and anxiety. To do this study, 35 sixth graders participated in the study, with 17 being placed in the experimental group and 18 being placed in the control group. According to the study's findings, students in the experimental group experienced significantly less writing anxiety than students in the control group. These findings show that the implementation of peer assessment reduced students' writing anxiety and increased their motivation for writing.

Zheng et al. (2021) also showed an interest in comparing the efficacy of self-assessment and peer assessment on oral skills. As per their mixed-method procedure, in a 16-week English public speaking course, two entire classrooms were randomly split into two groups, with 25 students placed in the self-assessment-initiated group and 26 students placed in the peer-assessment-initiated group. Both groups had to finish three formal English public presentations after two distinct sets of formative practice. The learners in the self-assessment-initiated group had substantially lower public speaking anxiety than those in the peer-assessment-initiated group, according to both quantitative and qualitative findings, whereas the learners in the peer-assessment-initiated group demonstrated significantly superior public speaking performance. The results also revealed that while peer-assessment-initiated formative practice may have helped learners enhance their correct use of the English language, the self-assessment-initiated formative practice helped learners pay more attention to their delivery and anxiety levels in public speaking.

Another applied study that investigated the impact of peer assessment on speaking skills is that of Hasnani and Mubarak (2020). Their study sought to determine whether or not peer assessment enhances the speaking skills of ESP students. To test if ESP students' speaking skills were improving, the researchers conducted an experimental study utilizing a peer assessment method. Thirtysecond-semester Teknik Laboratorium Medis students from Universitas Megarezky participated in the study. Data from the study's assessment, interviews, observations, and field notes were evaluated. Prior to implementing peer assessment, the average student score was 67.3; however, after using peer assessment, the average student score increased to 82. After employing peer assessment, the post-test revealed an improvement in speaking skills. Thus, peer evaluation proved to be effective in helping ESP students perform better when speaking. This study concluded that using peer assessment helps ESP students develop their oral skills.

In another attempt, Meletiadou (2021) explored the impact of peer assessment on EFL students' writing performance. Their study used a pre-test-post-test quasi-experimental method and sought to understand how peer assessment affected the writing abilities of 200 Greek Cypriot EFL students. Over the course of a complete academic year, adolescents had two 90-min writing classes each week. After receiving training in peer assessment techniques, teachers were required to instruct their own students. APA rubric, which was also developed by the researcher but agreed upon by the students and

their teachers throughout the training sessions, was given to the students to utilize. By comparing the results of the pre- and post-tests, paired t tests were used to determine whether or not the students in the control group ($n = 100$ students and 10 instructors) and the experimental group ($n = 100$ students and 10 teachers) improved their writing performance. The results of the study suggested that peer assessment might have a somewhat good effect on pupils' writing abilities. Students' writing performance increased in five areas thanks to peer assessment: mechanics, organization, substance, focus, and vocabulary and language use. This study offers suggestions for PA implementation in high school EFL writing classes that allow teachers to enhance students' writing performance in response to the need for greater experimentation.

Another study that examined the effect of self and peer assessment on EFL learners' argumentative writing performance is that of Iraji et al. (2016). Their study sought to determine how Iranian intermediate EFL students' performance in argumentative writing was influenced by their own and their peers' evaluations. Based on the results of the Oxford Quick Placement Test and an argumentative essay that functioned as the study's pre-test, 36 intermediate EFL students were chosen and homogenized for this purpose. They were subsequently divided into control and experimental groups and allocated at random to receive various therapies. Participants' compositions in the control group were evaluated by a traditional teacher, whereas the experimental group's works were evaluated by their peers and themselves. The utilization of self- and peer-assessments had a considerable impact on the learners' writing abilities, according to the post-test results. Based on the results of the study, it was determined that using alternative assessments for Iranian EFL students may assist them to overcome some of their challenges with argumentative writing. The findings of this study have significant ramifications for ELT stakeholders, including teachers, students, and other stakeholders. These alternative tests can be used as a teaching tool to help pupils become less anxious and develop their persuasive writing abilities.

In another research, Ritonga et al. (2022) explored the contribution of peer evaluation to text comprehension, reading motivation, and vocabulary learning in an EFL context. They used a convenience selection method to choose 60 Iranian EFL students at the intermediate level, ranging in age from 18 to 26, and divide them into two groups: the experimental group and the control group. Thereafter, the subjects were pretested on the constructs of the study. With the assistance of the teacher, the students in the experimental group evaluated the work of their other classmates. After each test in the control class, the students evaluated their own performance with the teacher's assistance. The post-tests for vocabulary, reading motivation, and reading comprehension were administered to both groups following a 15-session course of treatment. The results of the one-way ANCOVA tests showed that on the three post-tests of reading comprehension, reading motivation, and vocabulary learning, the experimental group outperformed the control group. In fact, the outcomes showed that peer assessment had a good impact on Iranian EFL students' text comprehension, motivation to read, and vocabulary learning. These L2 researchers concluded that peer assessment is an effective way for enhancing language learning among EFL students.

Although the studies outlined above directly pointed to the efficacy of peer feedback and cooperative/scaffolded learning, the literature on their usefulness in developing

vocabulary learning and oral skills of EFL learners is quite thin, and most experimental studies have mainly focused on the contribution of peer assessment and cooperative/scaffolded learning to reading comprehension and writing skills (Ritonga et al., 2022). To put it differently, there is a dearth of research on the effects of peer assessment on the growth of EFL learners' oral skills and lexical development. Therefore, by conducting this study, the researchers hope to close the knowledge gap and provide answers to the following two questions. Thus, the objectives of this study are twofold. The first objective of this study is to check whether implementing peer assessment/scaffolded learning has any distinguishing effect on learners' oral skills. The second objective of the study is to examine the potential contribution of peer assessment/scaffolded learning on EFL learners' vocabulary retention and recall.

> Research question 1: *To what extent can peer assessment/scaffolded learning ameliorate EFL learners' oral skills?*
> Research question 2: *To what extent can peer assessment/scaffolded learning facilitate vocabulary retention and recall among EFL learners?*

These research questions will lead to the following null hypotheses:

> Null hypothesis 1: Peer assessment/scaffolded learning cannot ameliorate learners' oral skills among EFL learners.
> Null hypothesis 2: Peer assessment/scaffolded learning cannot facilitate vocabulary retention and recall among EFL learners.

The studies reviewed above established the efficacy of peer feedback and cooperative/scaffolded learning. Notwithstanding, there is little empirical research on how these practical strategies can help EFL learners improve their both speaking skills and vocabulary learning. Instead, most experimental studies have primarily concentrated on how peer assessment and cooperative/scaffolded learning can improve reading comprehension and writing skills (Ritonga et al., 2022). To put it another way, little study has been done on the impact of peer assessment on the development of lexical and oral abilities in EFL learners. It is for this reason that this study seems to be an innovation, and attempts to close the current knowledge gap, and add to the literature on how these attempts can actually provide insights for policy-makers, materials developers, language teachers, etc. on how to implement the potential results of this study in their professional work.

Methodology

Design

In order to conduct this study, a quasi-experimental design was used because the researchers were unable to randomly choose the subjects. In this study, there were two groups: an experimental and a control group. In this study, in each research question, a two-level categorical independent variable (peer assessment and scaffolded learning vs. no peer assessment and scaffolded learning) and a dependent variable (the scores on a speaking test in the first research question, and the scores on a vocabulary test in the second research question) will be employed. It should also be mentioned that vocabulary

knowledge will be tested on three occasions (pre-test, post-test, and delayed post-test). Thus, the variable has three levels.

Participants

Cluster random sampling was used to select the participants in this study. From among language institutes in Khuzestan, Iran, 37 English as a Foreign Language (EFL) students from two preassembled general English classrooms at one of the language schools participated in this study. There were 120 language learners in the institute, but through an Oxford Quick Placement Test, 37 subjects with lower-intermediate, and 5 learners with an intermediate level of English proficiency were selected. They ranged in age from 15 to 20 years and had Farsi as their L1. Before the study, none of the participants had ever studied abroad and had little chance to communicate with others in English outside of the classroom. As students, they had taken three hours of English each week at school and another four at the language school. The experimental group ($N = 25$) and the control group ($N = 17$) were randomly assigned to each group. The experimental group was further divided into five groups of five, but the control group was not divided into groups. One intermediate learner was selected as the more proficient member (i.e., mediator) in each group and served as the group's head. The mediator further was responsible for providing other members of the group with appropriate assistance (i.e., mediation). It should not be forgotten that the scores of these mediators were not taken into account neither in the pre-test nor in the post-test as the objective of this research was to find an appropriate intervention for learners with a lower-intermediate level of language proficiency. It is also worth noting that through passive deception the researchers informed participants of the research objectives and their signed consent was obtained. For the participants under 18, parental signed consent was obtained.

Instruments

At the start of the study, the researchers evaluated the participants' vocabulary knowledge using an instructor-made lexical test. Based on the Vocabulary for the High School Student Book (Levine et al., 2004), a teacher-made vocabulary test with 20 multiple-choice questions was produced and administered as the study's pre-test. Thus, this instrument was adopted for the purpose of the study. The legitimacy of the test's face and content was checked by three validators. Additionally, the test was also administered to a group of language teachers who had the construct (i.e., they already knew the lexical items). The difference between these teachers' knowledge of the lexical items and those of the participants of the study was significant in a statistical sense ($p < 0.05$). This known group technique (Ary et al., 2019) also verifies the construct validity of the instrument. The KR-21 findings revealed that the vocabulary test's reliability index came out to be 0.81 indicating the reliability of the instrument. The students in the two groups took a similar test with the same number of items as the vocabulary post-test following the intervention. The validity and reliability ($r = 0.78$) of this post-test were also corroborated using the above-mentioned procedures. Additionally, using the same procedures, a similar vocabulary test with different items was administered two weeks after the treatment as the delayed post-test. The validity and reliability ($r = 0.81$) of the instrument were also determined.

Additionally, to check learners' oral proficiency an oral narrative task was also employed in this study. The task required the students to recount a story aloud following a series of Heaton photographs (Heaton, 1975). In groups of four, students in the experimental group were asked to narrate a tale based on the comic strip "A Surprise." Each student in this activity had two minutes to look at the photographs before being asked to tell a tale in response. They were required to tell the tale using verb tenses in regular type in simple past and simple present tense. In four different sessions, a total of four narrative tasks were used, and learners received the post-test right after the fourth task. The extent to which the subjects were able to demonstrate their command of the form in their production fluently, accurately, and with complexity constituted their speaking score in this study. The speaking score could go as low as 0 and as high as 20 in this study. To be more specific, for evaluation, a set of criteria developed from the TOEFL iBT Speaking Rubrics and the CEFR's 'Qualitative Aspects of Language Use' were employed. Based on the evaluation criteria, which included task completion, fluency, organization, and language use, scores between 0 and 20 were generated. A sample of English literature B.A. holders piloted the test to validate it. Two evaluators with a combined total of more than 15 years of experience instructing English at the university level examined each student's performance across the task to confirm its reliability. The task's average score was calculated and taken into account. In addition, to verify the inter-rater reliability, which was high for both pre- and post-test scores, Pearson's correlation coefficient was determined for the pre- and post-test scores (with 0.87 in time 1 to 0.89 in time 2). It goes without saying that because this second instrument was based on Heaton's (Heaton, 1975) photographs, this instrument was adapted for the purpose of this study.

Procedure

For the researchers to come up with homogenous groups, the subjects were administered an OQPT first. Only students with lower-intermediate proficiency were chosen for this study's participation as a consequence of the test. Since this study was conducted in a classroom context, other students whose scores varied from the rest were not taken into account. The two groupings were then made up of two classes. Peer evaluation, group work, and scaffolding were given to an experimental group, but neither of these activities was given to the control group. That is, the control group only received a traditional model of language teaching using a presentation, practice, and production approach. The entire experimental class was split into five groups of four students. For the oral narration activities, learners in the group were given six photographs and given two minutes to look at them. They were then asked to narrate a portion of the task. Following the completion of each learner's own narration in each group, peers from that group or other groups gave feedback to the learners by evaluating their narration. Peers were asked to evaluate the performance of their friends in terms of both language use and content. Each group went through this process until the entire class had completed the assignment. The process was repeated a second and third time to ensure the effectiveness of peer assessment on oral skills. This treatment lasted three sessions, and attempts were made to make learners more fluent in speaking by implementing Vygotskian ideas in the classroom.

In addition to the oral narrative task, the first chapter of Vocabulary for the High School Student by Levine et al. (2004), was instructed to the subjects in four sessions. The first chapter of the book is divided into six units. Thus, this part of the treatment lasted six sessions. That is each session was devoted to one unit of the first chapter of the book. One exam of ten vocabulary items was administered to both groups at the start of each session. Under the guidance of the teacher, the students in the experimental class evaluated the work of their peers. That is, peers looked at the responses made by their friends and provided guidelines on how their friends can correct their erroneous responses. After each test, the students in the control group reviewed their own results with the aid of the teacher and discussed their strengths and flaws. Thus, the experimental group was exposed to peer assessment, while the control group was not. After the treatment sessions, the vocabulary post-tests were given to both groups, and the results of the pretests and post-tests were compared and contrasted.

Method of data analysis

In each of the research questions raised above, there is a two-level categorical independent variable (peer assessment and scaffolded learning vs. no peer assessment and scaffolded learning) and a dependent variable (the scores on a speaking test in the first research question, and the scores on a vocabulary test in the second research question). Additionally, following recommendations made by Ary et al. (2019), as the means of two groups, namely the experimental group and the control group need to be compared for each of the research questions, two independent samples *t* tests must be conducted. Before conducting the *t* tests, for each of the questions a one-sample Kolmogorov-Smirnov test will be conducted to examine whether data are normally distributed or not. If the results of the test show that the data are normally distributed, the independent samples t-test will be conducted, otherwise, the Mann-Whitney *U* test as the non-parametric version of the *t* test (Ary et al., 2019) will be run. It is also crystal clear that the data will be examined through SPSS software.

Results

The contribution of peer assessment plus scaffolding to speaking skills

As mentioned in the previous section, as the means of two groups require to be compared, an independent-sample *t* test or a Mann-Whitney *U* test must be conducted (Ary et al., 2019). At this stage, through a one-sample Kolmogorov-Smirnov test, it is checked whether the data are normally distributed (Rezai, 2015).

As the results of Table 1 reveal, before the commencement of the treatment, data are normally distributed ($p > 0.05$). Thus, an essential assumption of independent samples *t* test is met, hence room for its conduction (Table 2).

The above-presented table shows the mean and standard deviation for both groups prior to and after implementing the treatment. It shows that the mean for both groups before implementing the treatment is almost the same ($M = 4.75$ for the experimental group with a 2.42 standard deviation and $M = 4.64$ with a 1.69 standard deviation for the control group). It implies that their knowledge was the same before the intervention. However, the experimental subjects performed much better compared to their control

Table 1 One-sample Kolmogorov-Smirnov test

	Group	Pre-test score
N	37	37
Normal parameters[a,b]		
Mean	1.45	4.70
Std. deviation	.50	2.09
Most extreme differences		
Absolute	.35	.14
Positive	.35	.14
Negative	− .31	− .10
Kolmogorov-Smirnov Z	2.18	.88
Asymp. Sig. (2-tailed)	.00	.41

[a] Test distribution is normal

[b] Calculated from data

Table 2 Group statistics

	Group	*N*	Mean	Std. deviation	Std. error mean
Pre-test score	Experimental	20	4.75	2.42	.54
	Control	17	4.64	1.69	.41
Post-test score	Experimental	20	11.50	4.65	1.03
	Control	17	6.47	2.34	.56

Table 3 Independent samples test

	Levene's test for equality of variances		*t* test for equality of means							
	F	Sig.	*t*	df	Sig. (2-tailed)	Mean difference	Std. error difference	95% confidence interval of the difference		
								Lower	Upper
Pre-test score									
Equal variances assumed	2.77	.10	.14	35	.88	.10	.70	1.31	1.52
Equal variances not assumed			.15	33.83	.88	.10	.68	1.27	1.48
Post-test score									
Equal variances assumed	7.74	.12	4.03	35	.00	5.02	1.24	2.50	7.55
Equal variances not assumed			4.24	29.00	.00	5.02	1.18	2.60	7.45

group counterparts ($M = 11.50$ for the experimental subjects with a 4.65 standard deviation and $M = 6.47$ for the control group with a 2.34 standard deviation).

Table 3 indicates that the homogeneity assumption underlying the independent samples *t* test is assumed as the sig. value both before and after the instruction is above 0.05 ($p > 0.05$).

Table 4 One-sample Kolmogorov-Smirnov test

	Group	Pre-test score
N	37	37
Normal parameters[a,b]		
Mean	1.45	3.59
Std. deviation	.50	1.58
Most extreme differences		
Absolute	.35	.13
Positive	.35	.13
Negative	− .31	− .13
Kolmogorov-Smirnov *Z*	2.18	.82
Asymp. sig. (2-tailed)	.00	.50

[a] Test distribution is normal

[b] Calculated from data

Table 5 Descriptive statistics

	Mean	Std. deviation	*N*
Pre-test score	3.59	1.58	37
Post-test score	6.51	3.79	37
Delayed post-test score	6.64	3.58	37

On the whole, an independent samples *t* test was run to see whether the learners who receive peer assessment and scaffolded learning perform better on an oral narrative task or not. An inspection of the data using a Kolmogorov-Smirnov test and a Levene's test showed no violation in normality and homogeneity assumptions of the *t* test. Although there was no significant difference between the subjects before the treatment ($p > 0.05$), the difference between the groups after the intervention was statistically significant. That is, the experimental group ($M = 11.5$, SD $= 4.65$) outperformed the control group ($M = 6.47$, SD $= 2.34$) on the post-test ($p < 0.05$, $t = 4.03$, df $= 35$). Moreover, the magnitude of the difference between the two groups, using Cohen's formula (1998), was quite large (eta squared $= 0.77$).

The contribution of peer assessment plus scaffolding to lexical gain

Based on the second research question, the researchers want to find out whether scaffolded peer assessment can affect vocabulary retention and recall. It is for this reason that learners need to run a repeated measure ANOVA (Ary et al., 2019; Rezai, 2015). First of all, there is a need to ensure that data are normally distributed. Thus, before running the repeated measures ANOVA, a one-sample Kolmogorv-Smirnov test needs to be carried out.

Data are normally distributed ($p > 0.05$) with relation to vocabulary knowledge prior to the start of the treatment, as shown by the results of Table 4. Thus, a prerequisite for the ANOVA is satisfied, allowing for its application.

Table 5 shows that for the pre-test score the mean for all participants was 3.59 with a 1.58 standard deviation ($M = 3.59$, SD $= 1.58$). The table further shows that the mean in

the immediate post-test boomed ($M = 6.51$, SD $= 3.79$), and the same holds true for the delayed post-test ($M = 6.64$, SD $= 3.58$).

As Table 6 reveals, there is a significant mean difference between the experimental group and the control group at $p > 0.05$ ($M = 3.54$, SD $= .61$). That is the experimental group significantly outperformed the control condition.

As shown in Table 7, the value for Wilks' lambda is 0.58 with a probability value of 0.001 indicating that there is a statistically significant effect for time. The effect size of 0.41 demonstrates a large effect size.

According to Table 8, the mean difference between time 1 (pre-test) and time 2 and 3 (immediate post-test and delayed post-test, respectively) is significant at 0.05. All in all, the above table using the Bonferroni adjustment test indicates that at the onset, the students were not familiar with the words, but after scaffolded peer assessment, their knowledge of vocabulary rose and this gain was durable over time.

All in all, for the second research question, a one-way repeated measures ANOVA was conducted to examine the effect of scaffolded cooperative learning on three occasions on EFL learners' vocabulary retention and recall. The descriptive results (mean and standard deviation) a statistically significant effect for time. Wilk's lambda $= 0.58$, F (2,

Table 6 Pairwise comparisons

	(J) Group	Mean difference (I-J)	Std. error	Sig.	95% confidence interval for difference	
					Lower bound	Upper bound
Experimental	Control	3.51	.61	.00	2.26	4.76
Control	Experimental	− 3.51	.61	.00	− 4.76	− 2.26

Table 7 Multivariate tests

Effect	Value	*F*	Hypothesis df	Error df	Sig.	Partial eta squared
Time						
Pillai's trace	.41	12.48	2.00	35.00	.000	.41
Wilks' lambda	.58	12.48	2.00	35.00	.000	.41
Hotelling's trace	.71	12.48	2.00	35.00	.000	.41
Roy's largest root	.71	12.48	2.00	35.00	.000	.41

Table 8 Pairwise comparisons

(I) Time	(J) Time	Mean difference (I-J)	Std. error	Sig.	95% confidence interval for difference	
					Lower bound	Upper bound
1	2	− 2.91	.63	.00	− 4.50	− 1.33
	3	− 3.05	.60	.00	− 4.56	− 1.54
2	1	2.91	.63	.00	1.33	4.50
	3	− .13	.25	0.97	− .76	.49
3	1	3.05	.60	.00	1.54	4.56
	2	.13	.25	0.97	− .49	.76

35) = 12.48, $p = 0.001$, multivariate eta squared = 0.41 indicating a large effect size. The pairwise comparison of the results showed that the scores increased from time 1 to time 2 and from time 1 to time 3 ($p = 0.001$), while the scores did not change from time 2 to time 3 ($p > 0.05$).

To sum up, regarding the potential contribution of scaffolded peer assessment on speaking skills, an independent samples t test was carried out. The results showed that the experimental group significantly outperformed the control group in the post-test, hence the efficacy of collaborative learning. Additionally, regarding the second research question, a one-way repeated measures ANOVA was conducted to examine the efficacy of scaffolded peer assessment on vocabulary retention and recall. The findings revealed that not only can the treatment result in lexical gain, but also its results are durable over time. That said, the results of the performed statistical analyzes point to the efficacy of peer assessment plus scaffolded learning in promoting both oral skills and lexical gains among EFL learners. For both dependent variables, learners who had experienced such training outperformed their peers who had not experienced scaffolded learning.

Discussion

The current study sought to investigate the effect of peer assessment interwoven in a group activity where students had the opportunity to obtain peer scaffolding in oral outputs as well as in vocabulary learning. The study's conclusions corroborated the value of peer evaluation and group work in helping learners improve both their speaking and lexical gain. Thus, it can be argued with full confidence that scaffolded peer assessment was successfully implemented in the current study.

The gained results are in line with Zheng et al. (2021) who showed that peer-assessment-initiated formative practice helped learners enhance their correct use of the English language. Additionally, our findings are following Hasnani and Mubarak (2020) who investigated the impact of peer assessment on speaking skills. Their study results indicated that using peer assessment helped ESP students enhance their oral skills. Moreover, the gained results in the present research are advocated by Ritonga et al. (2022) who explored the effects of peer evaluation on text comprehension, reading motivation, and vocabulary learning in an EFL context. Their outcomes revealed that the experimental group who was trained by peer-assessment outperformed the control group on the three post-tests of reading comprehension, reading motivation, and vocabulary learning.

Our findings concur with those of Movahedi and Aghajanzadeh Kiasi (2021) who confirmed the effectiveness of peer-evaluation in enhancing the writing abilities of Iranian intermediate EFL students. Furthermore, the findings of our study are supported by Salem Almahasneh and Abdul- Hamid (2019) who examined the effects of peer assessment training on writing performance among Arab EFL high school students. They discovered that students who took part in peer assessment training gained higher scores on the post-test of writing. Similar evidence supports our findings from Tunagür (2021), who investigated the effect of peer assessment application on sixth-grade students' writing motivation and anxiety. He concluded that peer assessment reduced students' writing anxiety and increased their writing motivation.

Researchers have recently argued more and more in favor of the use of evaluation to foster learning in academic practice (Wiliam, 2018). Peer evaluation is a fundamental

component of formative assessment theories because it is thought to give teachers or students fresh knowledge about the learning process, ameliorating later performance (Pellegrino et al., 2001). The findings of the present study support the idea that peer evaluation might be a useful instructional strategy for raising student achievement at least in language programs dealing with speaking skills and vocabulary learning. The findings imply that peer assessment can play a significant formative role in classrooms since it is more successful that no assessment at all. The results indicate that designing classroom activities in a way that makes use of peer assessment can be a useful strategy to encourage learning and maximize the use of instructional resources by allowing the instructor to concentrate on helping students with more challenging and complicated tasks. This shows, pragmatically, that classroom teachers can apply peer assessment in a variety of ways and customize the peer assessment design to the unique characteristics and limitations of their classroom environments (Double et al., 2020).

Although in line with a huge number of studies which had pointed to the positive effects of peer assessment, collaborative learning, and scaffolded learning on language productive skills (e.g., Hasnani & Mubarak, 2020; Hung et al., 2016; Imani, 2021; Nicol et al., 2014; Zheng et al., 2021), very few to no study had looked into the impact of scaffolded peer assessment on vocabulary learning (Ritonga et al., 2022). Thus, this research can be seen as an innovation in an absolutist sense for considering the potential impact of this treatment on a novel dependent variable and the extent to which the manipulated independent variable (scaffolded peer assessment) can take account of variation in the dependent variable (vocabulary learning).

One key element in an effective learner-centered program is active engagement which is widely acknowledged (White, 2009). The benefits of collaborative learning, when conversation partners and the teacher actively communicate with one another, include increased learner self-confidence, active participation, and motivation (White, 2009). According to Kaye (1989), group work is beneficial for active learning because it encourages both an equal quantity of turn-taking and more reflective and thought-out input.

Peer evaluation significantly changed how students perceived the function of the teacher. With peer assessment in place and the instructor's role reduced, the students eventually started to assess their own learning with the aid of their peers and their own capacity for critical evaluation, removing the need for the teacher to serve as the sole assessor or judge. The students began to take greater control of their own education in this way. As a result, the students' reliance on the teacher alone was significantly diminished. This is in line with the study conducted by Shen et al. (2020) who suggested that peer assessment significantly reduces sole reliance on the teacher as the mere source of knowledge to an inordinate degree and in this way promotes learners' autonomy.

By using peer-assessment, additional input and feedback can be provided for the students that help them develop their language learning. The students in the peer-assessment group received more suggestions and criticism, which enhanced their capacity for speaking and vocabulary. The pupils felt more in control of their language learning thanks to peer-assessment. Since they were forced to analyze both their own

and their peers' performances, peer assessment helped pupils become more aware of classroom discussions. In fact, peer-assessment can encourage cooperation among students which ultimately leads to language learning. The mentioned advantages of peer-assessment can be considered as reasons for the gained results of the current investigation.

The efficacy of this intervention lies in the fact that in a group setting, students are encouraged by their more experienced and skilled peers and receive tailored assistance accordingly. When learners are required to give a precise response to a question or to find a solution to a dilemma, their peers are available to provide appropriate assistance. When a group member provides an unsatisfactory response, the group's more advanced students can explain why the response is inadequate. In turn, this explanation can encourage cooperation among group members, which can promote a deeper understanding of the material through the clarification and/or elaboration that occurs as a result of this mediation. Additionally, scaffolded cooperation of this kind can actually reduce learners' stress and encourage them to work more collaboratively and jointly with their more proficient peers. In this way, scaffolded peer assessment can result both in intriguing lessons and the development of speaking skills as well as vocabulary knowledge. Through scaffolded cooperative learning, learners can widen their ZPD and in this way, they can regulate their own learning.

Through the use of peer assessment, students can engage in cooperative learning where they are eager to help and assess their peers and take ownership of their own language learning achievements. This can lead to improved social skills, better assessment, and more accurate feedback. Since learners help each other, better learning takes place because they care about the group and its members and desire to accomplish the same goal. This is in line with the social interdependence theory (Slavin, 2011). Accordingly, given that learners of the same age participate in cooperative activities and serve as role models for one another's ZPDs, which are more beneficial than having the students work alone, the findings of this research confirm the social constructivism idea put out by Vygotsky (Webb, 2008).

Although the present study directly points to the efficacy of scaffolded peer assessment, some points are in order. Similar to the observation by Colognesi et al. (2020), it was observed that the text's structure and coherence received the most positive feedback from peers. Along with grammar, they discussed verbal and nonverbal communication. Therefore, it appears that peer assessment is a method for immediately putting what was learned from their prior experience as language learners into practice. While the peers hardly mention the evolution of ideas, they do not mention the aim of communication at all. It is for these reasons that the researchers feel, to be most successful, peer assessment needs to be combined with teacher assessment so that language learners also focus on the features that their peers do not dwell on.

Whereas the findings of the study established the effectiveness of peer feedback on both developing speaking skills and lexical gain, it is not to assume that there is no flaw in peer evaluation. For example, according to Woolhouse (1999), it might be challenging for peer evaluators to assign fair grades. Besides, students experience emotional bias when they give their peers low grades, according to Oldfield and Macalpine's (Oldfield & Macalpine, 1995) study.

Conclusion, implications, limitations, and future directions

This current investigation with its focus on the effect of scaffolded peer assessment on speaking skills and lexical gain established an effective teaching strategy for language teachers. Studies from the past demonstrate that learners initially lack the ability to regulate and monitor their own learning, but if they are actively involved in making decisions, they will learn how to learn more effectively (Nikolov, 2006). In terms of direction, scaffolding can raise learning outcomes and learning quality. This study validated the significant positive impacts of scaffolding on learning outcomes for vocabulary and speaking abilities. Vygotsky (1978) stressed the value of supportive relationships and social contact for learning. From this angle, scaffolding is anticipated to significantly boost learners' performance and learning (Bonk & Cunningham, 1998). It is obvious that this research significantly supports the idea. Scaffolded training is advantageous, to put it simply.

All in all, the findings of this study established the efficacy of cooperative learning through peer assessment/scaffolded learning on both speaking skills and vocabulary retention and recall. Thus, it can be argued that implementing Vygostkian thinking in a practical domain is a better approach to developing both oral skills and lexical knowledge than traditional presentation-practice-production based on skill-acquisition theory. Peer assessment/scaffolded learning can help learners regulate their own learning and in this way, they can extend their ZPD. That is, learners' potential for extending what they learned from the teaching context to other situations can be extended if teachers implement Vygotskian thinking in their practice.

This scaffolded peer assessment study raises several implications for language instructors, administrators, syllabus designers, curriculum developers, language policy makers, etc. The first pedagogical implication of this study is that language teachers can implement such training in their classes to expand their learners' ZPD at least when they are concerned with developing speaking skills and enlarging their learners' vocabulary size as this intervention proved to be highly successful.

By utilizing the constructive criticism received from peers and self-assessments, teachers can enhance the support they provide for their pupils. Teachers may find it valuable to employ peer assessment as it is an effective instrument for the development of speaking skills and vocabulary knowledge. Teachers can promote students' autonomy in the classroom in addition to employing a variety of evaluations by having them engage in peer-assessment activities. EFL students can progressively understand what it takes to become self-directed learners as a result. Activities that involve peer evaluation are another method by that teachers can encourage cooperative learning. Students studying English as a foreign language may benefit from being familiar with various types of assessment in general and peer assessment in particular if they want to improve their language learning. Additionally, by using peer and self-assessment exercises, students may pinpoint precisely where they need aid and support so that they can require assistance from their teachers.

A third pedagogical implication of this study is that peer assessment/scaffolded learning can provide an intriguing, exciting, and almost stress-free context for learning the bolts and nuts of language. By implementing cooperative learning strategies through peer assessment/scaffolded learning, not only can learners improve their proficiency in the language, improve their speaking skills and increase their vocabulary

size, but also a stress-free context for learning will be created in which cooperation is encouraged rather than harmful competitiveness that hinders mental growth.

Another implication of this study is for the administrators. The principals of government-run schools and heads or chairs of language institutes can encourage their academic staff to take benefit of the findings of this study in their classroom to both to promote autonomy and enhance their learners' productive skills of speaking and vocabulary learning.

This study has also an implication for material developers. Curriculum planners and/or syllabus designers can also look in detail into the findings of this investigation and implement the issues raised in this study in their forthcoming materials. The findings can help syllabus designers have a clearer understanding of peer-assessment and how it may impact EFL learners' linguistic development. It is advised for syllabus designers to include a variety of assessment types in their curricula because the peer-assessment type was proven to have a positive impact on EFL learners' speaking and vocabulary learning. The findings of this study may also assist material designers in creating a variety of tasks and activities that are suitably suited to the speaking and vocabulary abilities of EFL students.

Despite the numerous positives of this study, the research is not flawless, so there is some room for further research. Because the study was constricted to lower-intermediate learners of English, at this stage, it is unclear whether such a treatment can be beneficial for learners with different levels of language proficiency. It is for this reason that the researchers suggest prospective researchers carry out further research to uncover the effect of this intervention in developing learners' ZPD across different levels of language proficiency. Another limitation of the research is that the subjects of the study were not divided into different age groups. It is recommended that future researchers conduct some studies to figure out if this intervention can also ameliorate the speaking skills and vocabulary learning of language learners. It is also suggested that researchers interested in such a topic extend the geographical limitation of this study and select learners with different L1 backgrounds, in another area, etc. to check the efficacy of this treatment. Additionally, to check whether the results of this study are long-lasting, longitudinal studies are needed, hence another room for further research.

Abbreviations

ZPD	Zone of proximal development
EFL	English as a foreign language
ELT	English language teaching
TOEFL	Test of English as a Foreign Language
CEFR	Common European Framework of Reference for Languages
OQPT	Oxford quick placement test
ZPD	Zone of proximal development

Author's contributions
The author had the responsibility of the whole paper. The author(s) read and approved the final manuscript.

Authors' information
Mina Homayouni got her MA in TEFL from Shahid Chamran University of Ahvaz, Ahvaz, Iran. She is now teaching IELTS at different IELTS academic centers.

References

Ahmadpour, L., & Yousefi, M. H. (2016). The role of mobile-assisted language learning on EFL Learners' development of writing accuracy, fluency, and complexity. *Journal of Modern Research in English Language Studies*, *3*(4), 105–118.

Aljaafreh, A., & Lantolf, J. P. (1994). Negative feedback as regulation and second language learning in the zone of proximal development. *The Modern Language Journal*, *78*(4), 465–483.

Ary, D., Jacobs, L. C., Sorensen, C. K., & Walker, D. (2019). *Introduction to research in education*, (10th ed.,). Wadsworth/Cengage Learning.

Babaii, E., Taghaddomi, S., & Pashmforoosh, R. (2016). Speaking self-assessment: Mismatches between learners' and teachers' criteria. *Language Testing*, *33*(3), 411–437.

Baniabdelrahman, A. (2010). The effect of the use of self-assessments' performance in reading comprehension in English. *The Electronic Journal for English as a Second Language*, *14*(2), 147–156.

Bolukbas, F., Keskin, F., & Polat, M. (2011). The effectiveness of cooperative learning on the reading comprehension skills in Turkish as a foreign language. *Turkish Online Journal of Educational Technology-TOJET*, *10*(4), 330–335.

Bonk, C. J., & Cunningham, D. J. (1998). Searching for learner-centered, constructivist, and sociocultural components of collaborative educational learning tools. In C. J. Bonk, & K. S. King (Eds.), *Electronic collaborators: Learner-centered technologies for literacy, apprenticeship, and discourse*, (pp. 25–50). Lawrence Erlbaum Associates Publishers.

Bryan, C., & Clegg, K. (Eds.) (2019). *Innovative assessment in higher education: A handbook for academic practitioners*. Routledge.

Bryant, D. A., & Carless, D. R. (2010). Peer assessment in a test-dominated setting: Empowering, boring or facilitating examination preparation? *Educational Research for Policy and Practice*, *9*(1), 3–15.

Cartney, P. (2010). Making changes to assessment methods in social work education: Focusing on process and outcome. *Social Work Education*, *29*(2), 137–151.

Chong, A. Y.-L., Chan, F. T. S., & Ooi, K.-B. (2012). Predicting consumer decisions to adopt mobile commerce: Cross country empirical examination between China and Malaysia. *Decision Support Systems*, *53*(1), 34–43.

Colognesi, S., Vassart, C., Blondeau, B., & Coertjens, L. (2020). Formative peer assessment to enhance primary school pupils' oral skills: Comparison of written feedback without discussion or oral feedback during a discussion. *Studies in Educational Evaluation*, *67*, 100917.

Diaz-Rico, L. T., & Weed, K. Z. (2002). *The Crosscultural, language, and academic development handbook: a complete K—12 reference guide*.

Doo, M. Y., Bonk, C., & Heo, H. (2020). A meta-analysis of scaffolding effects in online learning in higher education. *International Review of Research in Open and Distributed Learning*, *21*(3), 60–80.

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

Echevarria, J., Vogt, M. E., & Short, D. (2004). *Making content comprehensible for English language learners: The SIOP model*, (2nd ed.,). Pearson Education.

Elleman, A. M., & Oslund, E. L. (2019). Reading comprehension research: Implications for practice and policy. *Policy Insights from the Behavioral and Brain Sciences*, *6*(1), 3–11.

Ellis, R. (2015). *Understanding second language acquisition*, (2nd ed.,). Oxford University Press.

Falchikov, N. (2007). The place of peers in learning and assessment. In *Rethinking assessment in higher education*, (pp. 138–153). Routledge.

Fauzan, U. (2014). The Use of Improvisations Technique to Improve The Speaking Ability of EFL Students. *Dinamika Ilmu*, *14*(2), 264–287.

Fulcher, G., & Davidson, F. (2007). *Language testing and assessment*. Routledge.

Fyfe, E. R., & Rittle-Johnson, B. (2016). Feedback both helps and hinders learning: The causal role of prior knowledge. *Journal of Educational Psychology*, *108*(1), 82.

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.

Hasnani, H., & Mubarak, M. H. (2020). The use of peer-assessment in enhancing ESP students' speaking ability. *Ethical Lingua: Journal of Language Teaching and Literature*, *7*(1), 74–81.

Heaton, J. B. (1975). *Beginning composition through picture*. Pearson Education Limited.

Heshmat 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.

Hoffman, B. (2019). The influence of peer assessment training on assessment knowledge and reflective writing skill. *Journal of Applied Research in Higher Education*, *11*(4), 863–875.

Hung, Y.-J., Samuelson, B. L., & Chen, S. (2016). The relationships between peerand 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 International Publishing.

Hyland, F. (2000). ESL writers and feedback: Giving more autonomy to students. *Language Teaching Research*, *4*(1), 33–54.

Imani, S. (2021). 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.

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

Jiang, P., Namaziandost, E., Azizi, Z., & Razmi, M. H. (2022). Exploring the effects of online learning on EFL learners' motivation, anxiety, and attitudes during the COVID-19 pandemic: a focus on Iran. *Current Psychology*. https://doi.org/10.1007/s12144-022-04013-x.

Johnson, D. W., & Johnson, R. T. (1987). *Learning together and alone: cooperative, competitive, and individualistic learning*. Prentice-Hall, Inc.

Johnson, D. W., Johnson, R. T., & Smith, K. A. (2014). Cooperative learning: Improving university instruction by basing practice on validated theory. *Journal on Excellence in University Teaching*, *25*(4), 1–26.

Joordens, S., Pare, D. E., & Pruesse, K. (2009). PeerScholar: an evidence-based online peer assessment tool supporting critical thinking and clear communication. In *Proceedings of the 2009 International Conference on e-Learning*, (pp. 236–240).

Kaye, A. (1989). Computer-mediated communication and distance education. In R. Mason, & A. Kaye (Eds.), *Mindweave communication, computers and distance education*, (pp. 54–59). Pergamon Press.

Kearney, S. (2013). Improving engagement: the use of 'Authentic self-and peer-assessment for learning to enhance the student learning experience. *Assessment & Evaluation in Higher Education*, *38*(7), 875–891.

Khany, R., & Khosravian, F. (2014). Iranian EFL learners' vocabulary development through wikipedia. *English Language Teaching*, *7*(7), 57–67.

Khezrlou, S., Ellis, R., & Sadeghi, K. (2017). Effects of computer-assisted glosses on EFL learners' vocabulary acquisition and reading comprehension in three learning conditions. *System*, *65*, 104–116.

Kyndt, E., Raes, E., Lismont, B., Timmers, F., Cascallar, E., & Dochy, F. (2013). A meta-analysis of the effects of face-to-face cooperative learning. Do recent studies falsify or verify earlier findings? *Educational Research Review*, *10*, 133–149.

Levine, H., Levine, N., & Levine, R. (2004). *Vocabulary for the high school learner*, (4th ed.,). An Amsco Publication.

Liu, H., & Brantmeier, C. (2019). I know English: Self-assessment of foreign language reading and writing abilities among young Chinese learners of English. *System*, *80*, 60–72.

Liu, N. F., & Carless, D. (2006). Peer feedback: The learning element of peer assessment. *Teaching in Higher Education*, *11*(3), 279–290.

Long, M. H. (1996). The role of linguistic environment in second language acquisition. In W. C. Ritchie, & B. K. Bahtia (Eds.), *Handbook of second language acquisition*, (pp. 413–468). Academic.

McNamara, T. (2001). Language assessment as social practice: Challenges for research. *Language Testing*, *18*(4), 333–349.

Meletiadou, E. (2021). Exploring the impact of peer assessment on EFL students' writing performance. *IAFOR Journal of Education*, *9*(3), 77–95.

Mousavi, S. A. (2012). *An encyclopedic dictionary of language testing*. Rahnama Press.

Movahedi, N., & Aghajanzadeh Kiasi, G. (2021). Effects of teacher versus learner assessment activities on the Iranian intermediate EFL learners' writing ability. *International Journal of Research in English Education*, *6*(1), 49–63.

Namaziandost, E., Heydarnejad, T., & Rezai, A. (2022). Iranian EFL teachers' reflective teaching, emotion regulation, and immunity: Examining possible relationships. *Current Psychology*. https://doi.org/10.1007/s12144-022-03786-5.

Naserpour, A., & Zarei, A. A. (2021). Visually-mediated instruction of lexical collocations: The role of involvement load and task orientation. *Iranian Journal of Learning & Memory*, *3*(12), 39–50.

Nassaji, H., & Swain, M. (2000). A Vygotskian perspective on corrective feedback in L2: The effect of random versus negotiated help on the learning of English articles. *Language Awareness*, *9*(1), 34–51.

Nicol, D., Thomson, A., & Breslin, C. (2014). Rethinking feedback practices in higher education: A peer review perspective. *Assessment and Evaluation in Higher Education*, *39*, 102–122.

Nikolov, M. (2006). Test-taking strategies of 12- and 13-year-old Hungarian learners of EFL: Why whales have migraines. *Language Learning*, *56*, 1–51.

Nunan, D. (1999). *Second language teaching & learning*, (pp. 41042–42978). Heinle & Heinle Publishers.

Olaya, M. L., & González-González, G. M. E. (2020). Cooperative learning projects to foster reading skills. *GIST–Education and Learning Research Journal*, *21*, 119–139.

Oldfield, K. A., & Macalpine, J. M. K. (1995). Peer and self-assessment at tertiary level--An experiential report. *Assessment in Higher Education*, *20*(1), 125–132.

Panadero, E., Jonsson, A., & Strijbos, J. W. (2016). Scaffolding self-regulated learning through self-assessment and peer assessment: Guidelines for classroom implementation. In *Assessment for learning: Meeting the challenge of implementation*, (pp. 311–326). Springer.

Pantiwati, Y., & Husamah, H. (2017). Self and peer assessments in active learning model to increase metacognitive awareness and cognitive abilities. *International Journal of Instruction*, *10*(4), 185–202.

Patri, M. (2002). The influence of peer feedback on self-and peer-assessment of oral skills. *Language Testing*, *19*(2), 109–131.

Pellegrino, J. W., Chudowsky, N., & Glaser, R. (2001). *Knowing what students know: The science and design of educational assessment*. National Academy Press.

Race, P., Brown, S., & Smith, B. (2005). *500 tips on ASSESSMENT*, (2nd ed.,). RoutledgeFalmer.

Reinholz, D. (2016). The assessment cycle: A model for learning through peer assessment. *Assessment & Evaluation in Higher Education*, *41*(2), 301–315.

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.

Rezai, M. J. (2015). *ABC of SPSS for students of applied linguistics*. Yazd University Press.

Ritonga, M., Tazik, K., Omar, A., & Saberi 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*, *12*(1), 1–17.

Roschelle, J., & Teasley, S. D. (1995). The construction of shared knowledge in collaborative problem solving. In *Computer supported collaborative learning*, (pp. 69–97). Springer.

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.

Sawyer, R. K. (2006). Educating for innovation. *Thinking Skills and Creativity*, *1*(1), 41–48.

Shen, B., Bai, B., & Xue, W. (2020). The effects of peer assessment on learner autonomy: An empirical study in a Chinese college English writing class. *Studies in Educational Evaluation*, *64*, 100821.

Shih, C. M. (2010). The washback of the General English Proficiency Test on university policies: A Taiwan case study. *Language Assessment Quarterly*, *7*(3), 234–254.

Slavin, R. E. (1982). *Cooperative learning: Student teams. What research says to the teacher*. National Education Association Professional Library (Stock No. 1055-8-00; $2.00).

Slavin, R. E. (2011). *Student team learning: A practical guide to cooperative learning*, (3rd ed.,). National Education Association.

Sluijsmans, D. (1999). Creating a learning environment by using self-, peer- and co-assessment. *Learning Environments Research*, *1*, 293–319.

Syairofi, A., Mujahid, Z., Mustofa, M., Ubaidillah, M. F., & Namaziandost, E. (2022). Emancipating SLA findings to inform EFL textbooks: A look at Indonesian school english textbooks. *The Asia-Pacific Education Researcher*. https://doi.org/10.1007/s40299-022-00642-9.

Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of EDUCATIONAL RESEARCH*, *68*(3), 249–276.

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

Topping, K. J. (2009). Peer assessment. *Theory into Practice*, *48*(1), 20–27.

Tomas, W., & Palmer, E. (2018). Formative self-and peer assessment for improved student learning: the crucial factors of design, teacher participation and feedback. Assessment & Evaluation in Higher Education, 43(7), 1032–1047. https://doi.org/10.1080/02602938.2018.1427698.

Tunagür, M. (2021). The effect of peer assessment application on writing anxiety and writing motivation of 6th grade students. *International Journal of Education*, *1*(10), 96–105.

Ur, P. (1996). A Course in Language Teaching. Cambridge: Cambridge University Press.

Vadivel, B., Mathuranjali, M., & Khalil, N. R. (2021). Online teaching: Insufficient application of technology. In *Materials today: Proceedings*.

Van Zundert, M., Sluijsmans, D., & Van Merriënboer, J. (2010). Effective peer assessment processes: Research findings and future directions. *Learning and Instruction*, *20*(4), 270–279.

Verenikina, I. (2008). *Scaffolding and learning: Its role in nurturing new learners*.

Vygotsky, L. (1978). Interaction between learning and development. *Readings on the Development of Children*, *23*(3), 34–41.

Wajnryb, R. (1990). *Grammar dictation*. Oxford University Press.

Weaver, R.L., II, & Cottrell, H.W. (1988). Motivating students: Stimulating and sustaining student effort. *College Student Journal*, *22*(1), 22–32.

Webb, N. M. (2008). Learning in small groups. In T. L. Good (Ed.), *21st century education: A reference handbook*, (pp. 203–211). Sage.

White, E. (2009). Student perspectives of peer assessment for learning in a public speaking course. *Asian EFL Journal*, *33*, 1–36.

Wiliam, D. (2018). How can assessment support learning? A response to Wilson and Shepard, Penuel, and Pellegrino. *Educational Measurement: Issues and Practice*, *37*(1), 42–44.

Wood, D.J., Bruner, J.S., & Ross, G. (1976). The Role of Tutoring in Problem Solving. *Journal of Child Psychiatry and Psychology*, *17*, 89–100.

Woolhouse, M. (1999). Peer assessment: the participants' perception of two activities on a further education teacher education course. *Journal of Further and Higher Education*, *23*(2), 211–219.

Wright, V. (2018). Vygotsky and a global perspective on scaffolding in learning mathematics. In *Globalisation and education reforms*, (pp. 123–135). Springer.

Xu, L., Naserpour, A., Rezai, A., Namaziandost, E., & Azizi, Z. (2022). Exploring EFL learners' metaphorical conceptions of language learning: A multimodal analysis. *Journal of Psycholinguistic Research*. https://doi.org/10.1007/s10936-022-09842-2.

Yamashiro, A. D., & Johnson, J. (1997). Public speaking in EFL: Elements for course design. *The Language Teacher*, *21*(4), 13–17.

Yang, M., Badger, R., & Yu, Z. (2006). A comparative study of peer and teacher feedback in a Chinese EFL writing class. *Journal of Second Language Writing*, *15*(3), 179–200.

Zhao, H. (2010). Investigating learners' use and understanding of peer and teacher feedback on writing: A comparative study in a Chinese English writing classroom. *Assessing Writing*, *15*(1), 3–17.

Zheng, C., Wang, L., & Chai, C.S. (2021). Self-assessment first or peer-assessment first: effects of video-based formative practice on learners' English public speaking anxiety and performance. Computer Assisted Language Learning, 1–34 (In Press). https://doi.org/10.1080/09588221.2021.1946562.

Publisher's Note