Learner-oriented assessment (LOA) practice: the comparative study of self-assessment, peer assessment, and teacher assessment on EFL learners' writing complicity, accuracy, and fluency (CAF), speaking CAF, and attitude

Anwar Hammad Al-Rashidi[1], Muhammad Asif[2] , Masoumeh Ghasemi Vanani[3] and Amhara Aberash[4*]

*Correspondence:
amhara.aberash@gmail.com

[1] Counseling Psychology
from Imam Mohamed bin Saud
Islamic University, Department
of Psychology, College
of Education, Prince Sattam Bin
Abdulaziz University, Al-Kharj,
Saudi Arabia
[2] Department of English
and Literary Studies, University
of Management and Technology,
Lahore, Pakistan
[3] Islamic Azad University
of Najafabad, Najafabad, Iran
[4] Department of English
Language, Jimma University,
Jimma, Ethiopia

Abstract

As the effects of self-assessment, peer assessment, and teacher assessment on EFL learners' writing CAF and speaking CAF have not examined in Iran, this research compared their effectiveness on developing Iranian EFL students' writing CAF and speaking CAF. Moreover, this research examined the attitudes of EFL students towards the three types of assessments. To achieve these objectives, 75 Iranian intermediate EFL learners were selected based on a convenience sampling method and divided into three groups. Next, all groups were pretested on two skills of writing and speaking. Then, one group was taught writing and speaking skills by using teacher assessment, the second group was taught the same skills through using self-assessment; and the other group received the treatment by applying peer assessment. After an 18-session instruction, writing and speaking posttests were administered to the three groups, and also, three attitude questionnaires were distributed among them. The outcomes of one-way ANOVA tests revealed that the peer-assessment and self-assessment groups outflanked the teacher-assessment group on the writing and speaking posttests. The findings showed that the respondents of the three groups presented favorable attitudes towards the three sorts of assessments. It was concluded that the peer and self-assessment are useful techniques to develop EFL learners' writing and speaking skills. At the end of the study, a couple of implications and recommendations were enumerated.

Keywords: Attitude, Assessment, Peer assessment, Self-assessment, Teacher assessment, Writing CAF, Speaking CAF

Introduction

The assessment process is a crucial component of both learning and teaching. Without some sorts of assessment, it is not possible to confirm that instructional goals and objectives have been met. The results of assessments can have a significant impact on educators' and educational planners' evaluation of the effectiveness of current programs and

on their ability to identify effective approaches to advance the future course of action (Jafarigohar, 2017).

While teacher-centered assessment still predominates in most instructional situations, concerns about the shortcomings and potential drawbacks of traditional assessments have prepared the way for alternative assessments to slowly but steadily emerge. The development of alternative assessment, according to Al-Mahrooqi and Denman (2018), was primarily a response to the critiques of the traditional teacher-centered types of assessments. One of the most glaring examples of these shortcomings is the traditional assessment's focus on students' long-term retention of the materials they were taught at the expense of their capacity to express their creativity and exercise their autonomy in a variety of learning activities (Bourke & Mentis, 2011). In other words, the majority of conventional teacher-centered assessment methods do not push students'cognitive ability to go beyond the subject given and to develop sound solutions when confronted with novel issues.

The conventional teacher-oriented assessments that undervalue the role of the student have been contrasted with alternative assessment strategies that are learner-focused. Learner-centered approaches to assessment place an emphasis on the involvement of students in assessment processes and also their decision-making in learning and teaching (Coombe et al., 2007). Self- and peer assessment are two such cutting-edge learner-oriented assessment techniques. After finishing a learning activity, self-assessment is the process of evaluating one's performance and quantifying one's accomplishment (Richards & Schmidt, 2010). It is a crucial component of language learning and helps students develop their autonomy (Harris et al., 2015).

One of the assessments through which students can explore, develop, and determine their performance in relation to the course is self-assessment (Marzuki et al., 2020). Because it enables students to promote neutral learning and individual goal setting by raising their consciousness of their very personal learning, self-assessment is introduced to assist school students participate in class participation (Ma & Winke, 2019; Ünaldı, 2016). As part of the self-assessment process, college students determine the standards that apply to their work and make assessments of whether or not those standards might be met by their work (Ounis, 2017; Liu & Brantmeier, 2019; Tigchelaar, 2019). Additionally, self-assessment has been utilized frequently in English language testing to measure students' language proficiency in the four English language proficiency areas, including speaking, reading, writing, and listening (Abdul et al., 2018; Hung, 2019).

On the other hand, peer assessment is "a method for students to designate the level, worth, or quality of a work or performances of other equal status students" (Topping, 2009, p. 20). Peer assessment is a sort of formative assessment in which pupils receive feedback on their learning process from other pupils at the same level as themselves rather than from the teacher (Topping, 2009). It is a continual process that has been shown to empower learners with the ability and skills to examine and evaluate their peers' work, and ultimately of their own, in addition to enhancing students' strengths and addressing their deficiencies (Weaver & Esposto, 2011).

Peer assessment is a communicative method that language teachers use to actively include students in the learning process by asking them to evaluate the work of their peers in accordance with the teacher's standards. Peer assessment is a type of formative

evaluation in which the students themselves provide feedback on what they have learned (Fulcher & Davidson, 2007). This student-generated feedback is afterwards shared with peers and utilized to enhance performance in both oral and written tasks after students have received peer-assessment training. Peer assessment, according to Topping (2009), is a method of learning evaluation in which the learner evaluates the degree, worth, or caliber of a product or performance of other equal-status learners. Teachers encourage their students to participate actively in their education and the language creation of their peers by basing it on this hypothetical situation. Students are also given the opportunity to take on greater responsibility when making choices that may eventually affect their peers' work and achievement.

Using the mentioned assessments can be used to develop EFL learners' writing skill. Writing is an ability that reflects a person's mental and emotional state and fosters communication by giving one the chance to share experiences and views (Hidi & Boscolo, 2006). According to Topuzkanamış (2014), writing can be considered of as a language designed for the transfer of phenomena like feelings, thoughts, dreams, observations, experiences, and sensations, which are mental creations of the individual. The act of starting school lays the groundwork for writing. By actively participating in the individual's personal and social lives, this ability has a structure that simultaneously activates the high-level thinking abilities, affective, and cognitive structures of the person.

It can be said that the writing skill is developed more slowly and insufficiently than other talents due to the writing's complex structure; the difficulty to follow the feedback, correction, and evaluation phases; the intensity of the classroom environment; and teacher-caused issues (Karatay, 2013). Students become bored and reluctant since writing is such a tough process, which makes writing even more difficult (Kurudayıoğlu & Özay, 2010). Unpleasant feelings like dread and tension in writing may harm the creative abilities. Writing is significantly influenced by cognitive and emotional factors like attitude, motivation, anxiety, self-control, self-efficacy, and self-regulation (Ahıskalı, 2020).

Also, the mentioned assessments can be applied to improve EFL learners' speaking skill. Speaking is one of the various language skills that must be learned as part of the process of learning a language. For most people, speaking is equivalent to learning that particular language (Celce-Murcia, 2001). The training of speaking ability has also historically been one of the most difficult and significant duties; despite this, it is oftentimes disregarded and overlooked (Bora, 2012). The majority of EFL students, according to Richards and Renandya (2002), want to speak English. Additionally, they characterize English speaking as a multifaceted ability that enables speakers to accomplish a variety of goals in a variety of settings. They contend that in addition to having extensive syntactic and semantic knowledge, one needs to understand how a language is used in various social circumstances in order to become a proficient speaker of a foreign language. This is why learning to speak is challenging (Richards & Renandya, 2002).

Speaking has been described as a useful ability to convey message to others (Spratt et al., 2008). Speaking is often similar to knowing a language, according to Celce-Murcia (2001). Nunan (2003) disputes the idea that teaching and learning how to communicate in a second language is a difficult undertaking for both teachers and students, especially in situations when English is being taught and learnt as a foreign language. In relation to

the significance of speaking, Ellis (2004) asserts that acquiring the ability to express one-self in a language is the most frequent goal of language learning.

Using peer assessment and self-assessment can produce positive attitudes towards English language learning. According to Montano and Kasprzyk (2008), attitude is determined by one's beliefs of the results of engaging in a behavior and is weighed by assessments of those results. In this way, students who strongly feel that the activity will lead to positively valued results will have a good attitude towards the behavior. This mindset is likely to be detrimental to individuals who have strong convictions about the results that are not highly valued (Rashidi & Nazemi, 2015). According to Kara (2009), students who have favorable attitudes towards their learning process are more likely to engage with it and work hard to learn more. Additionally, it has been noted that these students are more eager to engage in emotional activity, solve issues, and learn skills and knowledge relevant to daily life.

Concerning the possible effects of self-assessment, peer assessment, and teacher assessment on language learning of EFL learners, the present research aimed at comparing the impacts of the mentioned assessments on enhancing Iranian EFL learners' writing CAF and speaking CAF. Furthermore, this research tried to examine the attitudes of EFL learners towards the three types of assessments.

The significance of this study lies in the fact that it includes three main assessments whose effects can greatly develop Iranian EFL learners' speaking and writing skills. Also, this study can be significant since it works on a new topic that has not been worked before. Not only it examines the effectiveness of the three assessments but also it investigates the attitudes of EFL learners towards the stated assessments. In addition, the findings of this research can add more new insights to the literature review in the testing and teaching domain.

Review of the literature

In the literature, assessment has been defined in several different ways. Among the many definitions, Linn and Miller (2005) definition of assessment of student learning as a systematic procedure for gathering data regarding students' advancement towards the learning goals stands out. They contend that a variety of methods can be used to gauge students' performance, including "conventional paper and pencil assessments, extended replies (essays), completion of authentic tasks, instructor observation, and student self-report" (p. 75).

Similar to this, Dhindsa et al. (2007) describe assessment as a crucial element of learning and teaching, "a systematic process of data gathering" regarding students' advancement (p. 1261). Teachers have a significant impact on how learners perceive evaluation, which affects how they learn (Schut et al., 2018; Watling & LaDonna, 2019; Zare Toofan et al., 2019). Since teaching involves human connection, the ways in which teachers engage with and relate to their students can have a significant impact on the type of learning environment they foster (Ramani et al., 2018; Telio et al., 2015).

Three categories of assessment exist: teacher, peer, and self-assessment. According to Dikel (2009), self-assessment is the judgment and evaluation that students make of their own learning. With the rising focus on learner-centered curriculum, needs analysis, and learner autonomy, self-assessment has become more and more popular

in recent years. Its potential utility as an instructional tool to support learning as well as a measuring tool has been a subject of much discussion (Butler & Lee, 2010). The desire to involve students as active participants in the learning process and to provide them the skills they need to be successful lifelong learners has led to an increase in interest in such methods (Jiang et al., 2022; Wood, 2009).

Self-assessment is connected to the students' involvement in evaluating their learning, particularly their accomplishments and the outcomes of their learning (Birjandi & Bolghari, 2015). Peer evaluation and cooperative learning go hand in hand (Fulcher & Davidson, 2007). The importance of collaboration is highlighted in this method, which encourages greater student contact. It is also one of the options that can combine instruction and assessment. Researchers concur that self- and peer-guiding assessment's principles have certain advantages (Brown & Hudson, 2012; Heshmat Ghahderijani et al., 2021).

Self-assessment is proposed to respond to opposition to the traditional sorts of assessments. The practice of accepting responsibility for learning and assessment enables the student to develop greater autonomy and self-governance (Chen, 2008). Self-assessment is connected to the students' involvement in evaluating their learning, particularly their accomplishments and the outcomes of their education (Birjandi & Bolghari, 2015). The use of self-assessment in the classroom helps students play a more active part in their education. This kind of evaluation is typically utilized for formative assessment goals to encourage the students to think about their learning process and outcomes (Harris & Brown, 2013). Additionally, Harris and McCann (1994) claimed that self-assessment is beneficial because it is quicker to ask students directly about their challenges.

The other type of assessment is peer assessment which is a type of group work that encourages social interaction and gives students a chance to support and learn from one another's actions and perspectives (Alias et al., 2015; Rezai et al., 2022). In order for this kind of collaboration to be successful, there needs to be a pre-established assessment criterion that students can use to organize their thoughts and make informed decisions about the work of their peers. Peer assessors will therefore participate in "cognitively demanding activities" during this collaborative effort, which will ultimately reinforce their performance and result in a deeper comprehension of the subject.

Peer assessment, according to Richards and Schmidt (2002), is "an alternate assessment technique in which students evaluate one another's progress using some checklists provided by their teachers" (p. 47). It is also a successful alternate assessment method (Richards & Schmidt, 2002). Additionally, it is one of the options that can combine training and assessment (Fulcher & Davidson, 2007).

Peer assessment is considered to be one of the main categories of option evaluation. Differentiated informative learning and instructive examination both emphasize the need of associate evaluation. Slavin (1997) suggested that associate appraisal was one of the pinnacle accomplishments in educational history. Peer assessment improves students' learning by giving them "a sense of ownership and responsibility, motivation, and impression of the students' own learning" (Saito & Fujita, 2009, p. 152). Shepard (2000) proposes peer assessment as one of the effective methods for classroom evaluation.

The students participate in peer assessment by ranking and evaluating one another's work. Through the act of exchanging fair and accurate feedback (Brown & Glasner, 2007) and making comparisons (Liu & Carless, 2006; Xu et al., 2022) with one another that are related to the outcome, it gives students the chance to acquire accountability and judgment skills. Peer assessment, as a source and tool that supports cooperative learning, enables the students to participate in evaluating one other's learning outputs and learning tasks, as well as in supporting and scaffolding each other's learning (Keith, 2005). According to their capacity for judgement, students give feedback to their peers. This has advantages, one of which is the development of student learning through cooperative learning (Brown & Glasner, 2007; Shafiee Rad et al., 2022).

In teacher assessment, the teachers are the ones who have a significant impact on students' learning because they evaluate that learning with the goal in mind (Wiggins & McTighe, 2005; Syairofi et al., 2022). According to Matsuno (2009), teacher assessment refers to evaluations in which the teacher examines and assesses the students' learning. The teacher keeps an eye on the pupils' actions, performances, and learning results and provides comments so that they can strengthen their areas of weakness. He suggested an alternative to psychometric testing by bringing up teachers' temporal characteristics and assessment standards.

Some empirical studies were carried out on the effectiveness of the mentioned assessments on developing English language learning. For instance, in a study on undergraduate students of China, Patri (2002) compared assessments of oral presentation skills made by teachers, peers, and the students themselves. After the students had some training sessions to become familiar with the assessment criteria, they were divided into two groups, one of which conducted self- and peer-evaluation while receiving peer feedback, and the other of which did not. By primarily using Pearson correlations to analyze the data, it was discovered that there was much more agreement between teachers' and peers' assessments when peer feedback was present than between instructors' and peers' assessments when peer feedback was present or absent.

A sample of Iranian English major students' paragraph writing abilities as well as the accuracy of their self- and peer ratings was examined by Birjandi & Siyyari, 2010 study. For this, pupils in two experimental groups composed eleven paragraphs over the course of eleven sessions, which they later evaluated either by themselves or their peers. The findings indicated that students' rating accuracy and writing ability both benefited from self- and peer assessment. Comparing the effects of self- and peer assessment on the participants' writing performance and rating accuracy, it was shown that peer assessment was more effective than self-assessment in improving the students' writing performance. Additionally, none of the assessment techniques performed better than the others in terms of improving the students' rating accuracy.

The impact of self-, peer-, and instructor assessments on Iranian undergraduate EFL students' course performance was examined by Abolfazli Khoonbi and Sadegh (2012). To do so, 82 university students were divided into four entire classes and assigned at random to one of the self-, peer, teacher, or control groups. The students' prior understanding of teaching methods was tested. The first experimental group underwent self-assessment activities, the second underwent peer-assessment tasks, and the third underwent teacher evaluation after obtaining pertinent instruction and training;

however, the control group underwent no assessment-related interventions. The posttest results showed that peer assessment performed better than self-assessment, teacher assessment, peer assessment, and the control groups.

Iranian EFL learners' speaking ability was the subject of an investigation by Ariafar & Fatemipour, 2013. Sixty Iranian EFL beginners participated in the study. The individuals in the experimental group got instruction and practiced self-assessment of their speaking abilities throughout the semester, but those in the control group had no such experiences. The findings showed that self-assessment exercises aided the study's participants in honing their speaking skills. Participants also expressed a favorable attitude towards self-assessment.

In another study, Heidarian (2016) examined the potential impact of self-assessment on EFL learners' writing skill. It was discovered that self-assessment significantly affected the writing ability of EFL learners. In the same line, Jafarigohar (2017) investigated how a cohort of Iranian English learners' motivation for writing and self-regulation were affected by three different evaluation procedures, including teacher, peer, and self-assessment. The researcher came to the conclusion that self-assessment was superior to the other two strategies in fostering writing motivation. The analysis of the coded think-aloud protocols and a one-way ANCOVA also supported the idea that alternative assessment, chiefly self-assessment, was superior in terms of encouraging the adoption of self-regulatory methods. Mazloomi and Khabiri (2018) investigated how dynamic self-assessment might affect EFL students' writing. The results demonstrated a considerable improvement in the writing abilities of EFL students utilizing dynamic self-assessment.

Using peer-assessment training, Salem Almahasneh and Abdul- Hamid (2019) looked at how writing performance among Arab EFL high school students was affected. There were 120 pupils between the ages of 15 and 16 in this research. The students came from two high schools in Malaysia that were Arab. Analytical marking scale was used to gather the data and evaluate the students' writing performance on the pretest and posttest. The findings of this study revealed a substantial difference between the experimental and control groups' writing abilities. According to the results, students who received peer-assessment training produced better writing drafts than those who received only traditional essay writing instruction and received no peer assessment.

Recently, Tunagür (2021) sought to determine whether the implementation of peer assessment has an impact on the writing motivation and anxiety of sixth-grade students. Students in sixth grade made up the study groups; 35 students participated in the study, with 17 assigned to the experimental group and 18 to the control group. The "writing anxiety scale" and the "writing motivation scale" were used to collect the study's data. A peer-assessment application was used for 6 weeks, during which the student writers' texts were assessed by their peers using a peer-assessment form. The research's findings led to the conclusion that the experimental group's students' writing anxiety dropped dramatically when compared to the control group's pupils. When the students' writing motivation was evaluated, it was discovered that the experimental group's pupils had greater writing motivation scores than the control group's students. According to these results, peer-assessment application decreased students' writing anxiety and boosted their enthusiasm to write.

The impacts of teacher versus student assessment activities on the writing proficiency of Iranian intermediate EFL students were examined by Movahedi & Aghajanzadeh Kiasi, 2021. Thirty intermediate students were chosen via the solutions placement test (SPT) to reply the questions of the study. They were then split into three groups of ten. Two experimental groups of students utilizing peer and self-assessment and one control group of students using teacher assessment were created. A pretest in writing was administered prior to the groups' eight sessions of therapy using peer, self, and instructor assessment techniques. After the instruction time, a writing posttest was given to all groups. The peer-assessment group scored the highest on the writing test, and it was discovered that there was a statistically significant difference between the effects of the teacher assessment, peer assessment, and self-assessment on the writing ability of Iranian intermediate EFL students. These findings were supported by the findings of the descriptive and inferential analyses.

Reviewing the literature shows that using peer assessment and self-assessment is more effective than the traditional assessments for English language learning. Most related studies were conducted on a single skill; very few studies were done on the effectiveness of the mentioned assessments on two or three skills and sub-kills simultaneously. Therefore, the current research compared the effects of the teacher-assessment, peer-assessment, and self-assessment assessments on boosting Iranian EFL learners' writing CAF and speaking CAF. Besides, this investigation inspected the attitudes of EFL learners towards the three sorts of assessments.

Based on the research objectives, the following questions were raised:

- **RQ1**: Does using peer assessment produce positive impacts on Iranian EFL learners' writing CAF?
- **RQ2:** Does using teacher assessment produce positive impacts on Iranian EFL learners' writing CAF?
- **RQ3**: Does using self-assessment produce positive impacts on Iranian EFL learners' writing CAF?
- **RQ4**. Does using peer assessment produce positive impacts on Iranian EFL learners' speaking CAF?
- **RQ5**: Does using teacher assessment produce positive impacts on Iranian EFL learners' speaking CAF?
- **RQ6**: Does using self-assessment produce positive impacts on Iranian EFL learners' speaking CAF?
- **RQ7**: Do Iranian EFL learners have positive attitudes towards self-assessment, teacher-assessment, and peer-assessment?

Methodology
Design of the study
Because we were unable to choose the participants at random for this investigation, we used a quasi-experimental approach. As a result, the participants in this study were chosen using a nonrandom sampling technique, and pre-test-treatment-post-test design was applied to collect the needed data. In this study, the peer-assessment and self-assessment

groups were regarded as the experimental groups, and the teacher-assessment group was considered as the control group. The speaking CAF, writing CAF, and attitude were the dependent variables, and teacher assessment, peer assessment, and self-assessment were the independent variables.

Participants

Seventy-five Iranian EFL students who were chosen from 107 EFL students according to their performance on the Oxford Quick Placement Test (OQPT) were participants in the study. The participants studied English at Ghoya English Institution, Isfahan, Iran. They were all intermediate-level students who ranged in age from 17 to 25. Using a convenience sampling technique, we chose the respondents and separated them into three groups at random: self-assessment, peer assessment, and teacher assessment. Each group contained 25 members. We could only work with female kids due to the gender segregation in the Islamic environment of Iran. It should be emphasized that when the participants signed the provided consent forms, the ethical standards were taken into account.

Instruments

The OQPT, which was created by the Oxford University Press, was the first tool used in this study to homogenize the subjects. It had 60 questions that assessed the pupils' reading comprehension, vocabulary, and grammar skills. The levels—elementary, pre-intermediate, intermediate, and advanced—at which their respondents were functioning would be better understood by the researchers as a result. The target population for the research was decided to be students who scored between one standard deviation (SD) above and below the mean.

The other tool employed in this survey was a pretest of speaking prepared by the researcher that included numerous items from the course book of the participants (i.e., top notch 3). The respondents had 2 to 3 min to discuss the topics, and their voices were recorded for the second rater (the respondents' speaking performances were evaluated by two raters). Applied linguistics professors of English have backed up the test's validity. Three university professors from Iran with more than 13 years of English teaching served as the validators. Additionally, using Pearson correlation analysis, the speaking test reliability was calculated as ($\alpha = .83$). It should be emphasized that this exam was used for both the research's speaking pretest and posttest.

A writing pre-intermediate pretest created by the researcher was the third and most crucial tool for acquiring the information required to answer the writing questions. It was based on the course book for the pupils (practical writer with readings). There were two subjects, and the students had to write an essay about one of them at random. The participant's composition on the chosen topic was to be written under the supervision of the researcher. The respondents were required to produce an essay of at least 150 words. To guarantee that the students complete the pretest independently, the researcher oversaw its administration in the classroom. Following the writing on the subject, all the essays were gathered and scored by two English teachers using some standards. When evaluating the students' writing ability, the raters took into account the grammatical accuracy, the sentences' meaning, and the length of each essay. Grammar, sentence

meaning, and length faults made by the pupils were all counted, after which they were given a grade. Two English specialists attested to the pre-test validity, and the inter-rater reliability of the test was determined using Pearson correlation analysis ($\alpha = 0.89$).

In the present investigation, a writing posttest was applied. The topics that were presented to the groups constituted the basis for the post-test. The students were required to write about one of the two subjects on the posttest. The pupils' compositions were graded by two raters. The posttest was given to participants to gauge how much their writing had improved as a result of the treatment. It should be mentioned that the validity of the post test was checked by two English experts, and its reliability was determined by utilizing Pearson correlation analysis ($\alpha = 0.87$).

Three attitude questionnaires were used as the last instrument in this study. One of them contained 20 items related to the self-assessment, and it was given to the self-assessment group to check their general perceptions about self-assessment. Another questionnaire had 20 items pertinent to the peer assessment, and it was administered to the peer-assessment group to examine their general perceptions about peer assessment. The last one included 20 items related to the teacher-assessment, and it was given to the teacher-assessment group to investigate their general attitudes towards teacher assessment. The researchers themselves made 20-point Likert items (highly disagree to highly agree) for each questionnaire and measured the reliability of each questionnaire (peer assessment $= \alpha = 0.81$, self-assessment $= \alpha = 0.84$, and teacher-assessment$= \alpha =.86$). The validities of the three questionnaires were proven by a group of English specialists.

The speaking checklist of Hughes (2003) was used to aid the raters score the participants' speech. The raters scored the participants' speaking CAF based on the criteria or scales in mentioned speaking checklist.

To analyze writing CAF, we mostly referred to the metrics created and applied by Wigglesworth and Storch (2009). While we added words per clause, Wigglesworth and Storch (2009) only utilized the ratio of clauses to T units and the percentage of dependent clauses of total clauses to measure complexity.

Collecting and analyzing the data

Three equal groups of 25 at the intermediate English proficiency level were selected to do this research. Next, all groups were pretested on two skills of writing and speaking. Then, one group was taught writing and speaking skills by using teacher-assessment; the second group was taught the same skills through using self-assessment, and the other group received the treatment by applying peer assessment. The peer-assessment group was assigned to five sub-groups of five, and peer assessment was used to measure their writing and speaking skills. Under the direction of the teacher, the students in the peer-evaluation group evaluated the work of their peers. With the teacher's assistance, the students in the self-assessment group analyzed their own performances following each test and made comments about their strengths and flaws. In the teacher-assessment group, the participants' performances on their exams were evaluated by their teacher. After an 18-session instruction, writing and speaking posttests were administered to the three groups, and also three attitude questionnaires were distributed among them.

The researchers used the aforementioned methods to get the necessary data, which they then examined to arrive at the final findings. First, they used the

Kolmogorov-Smirnov test to confirm that the data's distribution was normally distributed. Next, they used one-way ANOVA tests to evaluate the data from the pre- and posttests. Finally, they used a one-sample *t*-test to examine the data from the questionnaires.

Results of the study

Following evaluation of participants' performance on the pretests and posttests, the collected data were examined using SPSS software, version 23. The following tables show the obtained results (Table 1).

The mean, standard deviation, and standard error for each distinct group are included in the descriptive statistics provided by the preceding table (peer assessment, teacher assessment, and self-assessment). This table shows that the means of all groups on the pretests of speaking accuracy, fluency, and complexity were almost equal. As it can be observed under the mean column, the mean scores of all groups are between 12.12 and 13.73, implying that they performed similarly on their speaking pretests.

In Table 2, three one-way ANOVA tests are run to help the researchers determine the differences among the speaking pretests of the three groups. As clearly seen under the Sig. column, the sig. values are higher than the pre-determined Sig. (.05), accordingly, there is not a statistically significant difference among the groups' speaking pretests of accuracy, fluency, and complexity

In Table 3, the speaking posttests descriptive statistics (the mean, standard deviation, & standard error) of the three groups are depicted. The results in this table clearly indicate that the means of all groups on the posttests of speaking accuracy, fluency, and complexity are different; the two groups of peer assessment and self-assessment got higher scores on their speaking posttests compared to the teacher assessment group (Table 4).

The results of the ANOVA analysis are displayed in this table, along with whether there is a statistically significant difference in the mean scores between the groups. There is a statistically significant difference in the mean scores between the three groups, as

Table 1 Descriptive statistics of all groups on the pretests of speaking accuracy, fluency, and complexity

		N	Mean	Std. deviation	Std. error	95% confidence interval for mean		Minimum	Maximum
						Lower bound	Upper bound		
SFPRE	Teacher	25	12.29	2.07	0.42	11.41	13.16	8.00	16.00
	Self	25	12.80	1.78	0.35	12.08	13.53	9.00	16.00
	Peer	25	12.12	2.43	0.48	11.11	13.12	8.00	18.00
	Total	75	12.41	2.10	0.24	11.92	12.89	8.00	18.00
SAPRE	Teacher	25	13.25	1.59	0.32	12.57	13.92	10.00	16.00
	Self	25	13.73	1.40	0.27	13.16	14.29	11.00	16.00
	Peer	25	12.20	2.04	0.40	11.35	13.04	8.00	17.00
	Total	75	13.06	1.79	0.20	12.65	13.47	8.00	17.00
SCPRE	Teacher	25	12.70	1.65	0.33	12.00	13.40	10.00	16.00
	Self	25	13.26	1.45	0.28	12.68	13.85	11.00	16.00
	Peer	25	12.60	1.93	0.38	11.80	13.39	8.00	17.00
	Total	75	12.86	1.69	0.19	12.47	13.25	8.00	17.00

Table 2 Inferential statistics of all groups on the pretests of speaking accuracy, fluency, and complexity

		Sum of squares	df	Mean square	F	Sig.
SFPRE	Between groups	6.55	2	3.27	0.73	0.48
	Within groups	321.63	72	4.46		
	Total	328.18	74			
SAPRE	Between groups	31.05	2	15.52	5.38	.08
	Within groups	207.61	72	2.88		
	Total	238.66	74			
SCPRE	Between groups	6.59	2	3.29	1.15	0.32
	Within groups	206.07	72	2.86		
	Total	212.66	74			

Table 3 Descriptive statistics of all groups on the posttests of speaking accuracy, fluency, and complexity

		N	Mean	Std. deviation	Std. error	95% confidence interval for mean		Minimum	Maximum
						Lower bound	Upper bound		
SFPOST	Teacher	25	13.04	1.39	0.28	12.45	13.63	11.00	16.00
	Self	25	14.96	2.21	0.43	14.06	15.85	10.00	18.00
	Peer	25	16.96	1.36	0.27	16.39	17.52	14.00	19.00
	Total	75	15.01	2.32	0.26	14.47	15.54	10.00	19.00
SAPOST	Teacher	25	14.62	2.14	0.43	13.72	15.52	12.00	18.00
	Self	25	16.38	1.81	0.35	15.65	17.11	12.00	20.00
	Peer	25	17.08	1.35	0.27	16.52	17.63	15.00	20.00
	Total	75	16.05	2.04	0.23	15.58	16.52	12.00	20.00
SCPOST	Teacher	25	13.66	1.27	0.26	13.12	14.20	12.00	16.00
	Self	25	15.53	1.77	0.34	14.82	16.25	12.00	18.00
	Peer	25	16.44	1.19	0.23	15.94	16.93	14.00	19.00
	Total	75	15.24	1.82	0.21	14.81	15.66	12.00	19.00

Table 4 Inferential statistics of all groups on the posttests of speaking accuracy, fluency, and complexity

		Sum of squares	Df	Mean square	F	Sig.
SFPOST	Between groups	188.10	2	94.05	31.81	.00
	Within groups	212.88	72	2.95		
	Total	400.98	74			
SAPOST	Between groups	78.16	2	39.08	12.14	.00
	Within groups	231.61	72	3.21		
	Total	309.78	74			
SCPOST	Between groups	97.72	2	48.86	23.46	.00
	Within groups	149.95	72	2.08		
	Total	247.68	74			

can be seen from the fact that the significance values are all 0.01 (i.e., $p = .01$), which are lower than 0.05. To determine the exact differences among the speaking posttests of the three groups, a post hoc Scheffe test is done in the following table (Table 5).

As of now, the data indicate that the groups as a whole differ statistically significantly from one another. Which groups varied from one another is seen in the multiple comparisons table above. Although there are many alternative options, the Scheffe post hoc test is typically the one used for one-way ANOVA post hoc tests. One can observe that there are statistically significant differences among the posttests of all groups except the posttests of speaking accuracy of peer-assessment and self-assessment groups (Table 6).

The table above presents the descriptive statistics for the three groups of teacher assessment, peer assessment, and self-assessment, including the mean, standard deviation, and standard error. Based on the results, the mean scores of the three groups on all writing pretests are almost equal. By taking a look at the above table, we see that the mean scores of all groups are between 12.44 and 13.76, meaning that all groups conducted similarly on their writing pretests.

The results of the three one-way ANOVA tests in in Table 7 determine the differences among the writing pretests of the three groups. As clearly observed under the Sig. column, the sig. values are all above the Sig. (.05); therefore, there is not a statistically significant difference among the groups' writing pretests of accuracy, fluency, and complexity.

In Table 8, the writing posttests descriptive statistics (the mean, standard deviation, & standard error) of the three groups are shown. The outcomes in this table obviously show that the mean scores of all groups on the posttests of writing accuracy, fluency, and complexity are different; the two groups of peer-assessment and self-assessment got better scores on their writing posttests compared to the teacher assessment group (Table 9).

Table 5 Post hoc Scheffe test (all groups' posttests of speaking accuracy, fluency, and complexity)

Dependent variable	(I) groups	(J) groups	Mean difference (I-J)	Std. error	Sig.	95% confidence interval	
						Lower bound	Upper bound
SFPOST	Teacher	Self	−1.91	0.48	.00	−3.13	−0.70
		Peer	−3.91	0.49	.00	−5.14	−2.69
	Self	Teacher	1.91	0.48	.00	0.70	3.13
		Peer	−1.99	0.48	.00	−3.20	−0.79
	Peer	Teacher	3.91	0.49	.00	2.69	5.14
		Self	1.99	0.48	.00	0.79	3.20
SAPOST	Teacher	Self	−1.75	0.50	.00	−3.02	−0.49
		Peer	−2.45	0.51	.00	−3.73	−1.17
	Self	Teacher	1.75	0.50	.00	0.49	3.02
		Peer	−0.69	0.50	0.38	−1.95	0.56
	Peer	Teacher	2.45	0.51	.00	1.17	3.73
		Self	0.69	0.50	0.38	−0.56	1.95
SCPOST	Teacher	Self	−1.87	0.40	.00	−2.89	−0.85
		Peer	−2.77	0.41	.00	−3.80	−1.74
	Self	Teacher	1.87	0.40	.00	0.85	2.89
		Peer	−0.90	0.40	.00	−1.91	0.10
	Peer	Teacher	2.77	0.41	.00	1.74	3.80
		Self	0.90	0.40	.00	−0.10	1.91

Al-Rashidi *et al. Language Testing in Asia*   (2022) 12:59

Page 14 of 22

Table 6 Descriptive statistics of all groups on the pretests of writing accuracy, fluency, and complexity

		N	Mean	Std. deviation	Std. error	95% confidence interval for mean		Minimum	Maximum
						Lower bound	Upper bound		
WFPRE	Teacher	25	12.91	1.66	0.34	12.21	13.62	10.00	16.00
	Self	25	13.26	1.53	0.30	12.64	13.89	11.00	16.00
	Peer	25	12.44	1.87	0.37	11.66	13.21	8.00	17.00
	Total	75	12.88	1.70	0.19	12.48	13.27	8.00	17.00
WAPRE	Teacher	25	13.20	1.99	0.40	12.36	14.05	8.00	17.00
	Self	25	13.23	1.45	0.28	13.64	14.81	12.00	17.00
	Peer	25	13.32	1.93	0.38	12.52	14.11	11.00	17.00
	Total	75	13.60	1.83	0.21	13.17	14.02	8.00	17.00
WCPRE	Teacher	25	12.54	2.24	0.45	11.59	13.48	8.00	16.00
	Self	25	13.76	1.33	0.26	13.22	14.30	11.00	16.00
	Peer	25	12.80	1.73	0.34	12.08	13.51	10.00	17.00
	Total	75	13.05	1.85	0.21	12.627	13.47	8.00	17.00

Table 7 Inferential statistics of all groups on the pretests of writing accuracy, fluency, and complexity

		Sum of squares	df	Mean square	F	Sig.
WFPRE	Between groups	8.81	2	4.40	1.53	0.22
	Within groups	207.10	72	2.87		
	Total	215.92	74			
WAPRE	Between groups	15.98	2	7.99	2.45	.09
	Within groups	234.01	72	3.25		
	Total	250.00	74			
WCPRE	Between groups	21.21	2	10.60	3.28	.06
	Within groups	232.57	72	3.23		
	Total	253.78	74			

The outcomes of the ANOVA analysis are shown in this table, along with whether there is a statistically significant difference between the mean scores of the three groups. There is a statistically significant difference in the mean scores between the three groups, as can be seen from the fact that the significance values are all 0.01 (i.e., $p = .01$), which are less than 0.05. To determine the exact differences among the writing posttests of the three groups, a post hoc Scheffe test is conducted in the following table (Table 10).

Based on the results gained so far, one can understand that there are statistically significant differences between the groups as a whole. This table helps us to know which groups had different performances on the writing posttests. The results of Scheffe post hoc test indicate that the performances of the teacher assessment group on the three posttests of writing were different from the performances of the other two groups. The results demonstrate that there are not significant differences among the posttests of peer-assessment group and self-assessment group.

Table 8 Descriptive statistics of all groups on the posttests of writing accuracy, fluency, and complexity

		N	Mean	Std. deviation	Std. error	95% confidence interval for mean		Minimum	Maximum
						Lower bound	Upper bound		
WFPOST	Teacher	25	14.12	2.02	0.41	13.26	14.98	10.00	18.00
	Self	25	15.76	1.70	0.33	15.08	16.45	12.00	18.00
	Peer	25	16.56	1.55	0.31	15.91	17.20	13.00	20.00
	Total	75	15.50	2.01	0.23	15.04	15.97	10.00	20.00
WAPOST	Teacher	25	15.29	2.36	0.48	14.29	16.29	12.00	19.00
	Self	25	16.92	1.64	0.32	16.25	17.58	14.00	19.00
	Peer	25	16.84	1.24	0.24	16.32	17.35	15.00	20.00
	Total	75	16.37	1.92	0.22	15.92	16.81	12.00	20.00
WCPOST	Teacher	25	14.45	1.99	0.40	13.61	15.30	12.00	19.00
	Self	25	16.57	1.87	0.36	15.81	17.33	12.00	19.00
	Peer	25	16.12	1.09	0.21	15.66	16.57	13.00	18.00
	Total	75	15.74	1.91	0.22	15.30	16.18	12.00	19.00

Table 9 Inferential statistics of all groups on the posttests of writing accuracy, fluency, and complexity

		Sum of squares	Df	Mean square	F	Sig.
WFPOST	Between groups	75.34	2	37.67	12.03	.00
	Within groups	225.40	72	3.13		
	Total	300.74	74			
WAPOST	Between groups	41.38	2	20.69	6.36	.00
	Within groups	234.16	72	3.25		
	Total	275.54	74			
WCPOST	Between groups	61.24	2	30.62	10.55	.00
	Within groups	208.94	72	2.90		
	Total	270.18	74			

As shown in Table 11, the amount of statistic *T*-value is 50.23, 55.78 and 48.02, $df =$ 19, and the three *Sig.* values are lower than 0.05. This infers that Iranian students had positive attitudes towards the effectiveness of the mentioned assessments on their language learning development. In short, the findings show that peer-assessment and self-assessment groups outdid the teacher-assessment group on their posttests of speaking CAF and writing CAF. In fact, peer assessment and self-assessment are more effective than teacher assessment for learning English as foreign language. Also, the results indicate that participants held favorable attitudes towards all three assessments.

Discussion and conclusion

Respecting the impacts of teacher assessment, peer assessment, and self-assessment on boosting speaking CAF and writing CAF of Iranian EFL learners, the findings of the present study demonstrated that applying the peer assessment and self-assessment was more beneficial than the teacher assessment on the mentioned dependent variables

Table 10 Post hoc Scheffe test (all groups' posttests of writing accuracy, fluency, and complexity)

Dependent variable	(I) groups	(J) groups	Mean difference (I-J)	Std. error	Sig.	95% confidence interval	
						Lower bound	Upper bound
WFPOST	Teacher	Self	−1.64	0.50	.00	−2.89	−0.39
		Peer	−2.43	0.50	.00	−3.69	−1.17
	Self	Teacher	1.64	0.50	.00	0.39	2.89
		Peer	−0.79	0.49	0.28	−2.02	0.44
	Peer	Teacher	2.43	0.50	.00	1.17	3.69
		Self	0.79	0.49	0.28	−0.44	2.02
WAPOST	Teacher	Self	−1.63	0.51	.00	−2.90	−0.35
		Peer	−1.54	0.51	.01	−2.83	−0.26
	Self	Teacher	1.63	0.51	.00	0.35	2.90
		Peer	.08	0.50	0.98	−1.17	1.34
	Peer	Teacher	1.54	0.51	.01	0.26	2.83
		Self	−.08	0.50	0.98	−1.34	1.17
WCPOST	Teacher	Self	−2.11	0.48	.00	−3.32	−0.91
		Peer	−1.66	0.48	.00	−2.87	−0.44
	Self	Teacher	2.11	0.48	.00	0.91	3.32
		Peer	0.45	0.47	0.63	−0.73	1.64
	Peer	Teacher	1.66	0.48	.00	0.44	2.87
		Self	−0.45	0.47	0.63	−1.64	0.73

Table 11 One-sample test of the attitude questionnaires

	Test value = 0					
	T	df	Sig. (2-tailed)	Mean differences	95% confidence interval of the differences	
					Lower	Upper
Teacher	50.23	19	.04	3.25	3.98	4.25
Self	55.78	19	.02	5.02	4.32	4.69
Peer	48.02	19	.00	4.58	4.55	4.78

(speaking CAF and writing CAF). On their posttests, the teacher-assessment group and the peer-assessment and self-assessment groups actually performed in very different ways. Furthermore, the outcomes depicted that the respondents held positive attitudes towards utilizing the three types of assessments.

The results of Jafarigohar (2017) study, which looked at the influence of instructor, peer, and self-assessment on Iranian English learners' writing motivation and self-regulation, support our research findings. His research showed that self-evaluation promoted writing motivation more effectively than the other two strategies. Additionally, his findings supported the idea that peer and self-assessments are more effective than instructor evaluations at encouraging the use of self-regulation techniques.

The current study is further supported by Ariafar and Fatemipour (2013), who looked into how Iranian EFL learners' self-assessment of their speaking abilities affected their progress. Their findings showed that the study's participants were able to increase their

speaking skills, thanks to self-assessment exercises. Heidarian (2016) and Mazloomi and Khabiri (2018), who investigated the possible impact of self-assessment on writing of EFL learners, corroborate our findings. They discovered that self-evaluation significantly affected the writing of EFL learners.

Additionally, Salem Almahasneh and Abdul- Hamid (2019), who looked at the impact of peer-evaluation training on writing performance among Arab EFL high school students, support the findings of our study. According to their findings, students who participated in peer-assessment training performed better on the writing posttest. Tunagür (2021), who looked at whether peer-assessment application has an impact on writing anxiety and writing motivation of sixth-grade pupils, supports our findings in a similar manner. The use of peer assessment, he concluded, decreased students' writing anxiety and improved their motivation to write.

The results of the current study also concur with those of other researchers, such as Patri (2002), whose investigation confirmed the beneficial impacts of peer and self-evaluations on college students of Chinese oral presentation skills. Furthermore, the results of Abolfazli Khoonbi and Sadegh (2012), who corroborated the impacts of self- and peer assessment on Iranian undergraduate EFL students' course achievement, are consistent with our own findings.

Our results also agree with those of Movahedi and Aghajanzadeh Kiasi (2021), who supported the value of peer and self-evaluation in improving Iranian intermediate EFL students' writing skills. The current study's findings also align with those of Birjandi and Siyyari (2010), who found that writing performance among EFL students was enhanced by self- and peer evaluations. Additionally, Imani (2022) compared the effects of self-assessment and peer assessment on reflective and impulsive EFL learners' speaking skill, and his findings support our own. According to his findings of a two-way ANOVA, speaking skill was affected by self- and peer evaluation equally in impulsive and reflective learners.

The cooperation that peer evaluation fostered among the students may be one reason for the peer-assessment group's improved performances as compared to the teacher assessment group. Peer assessment can foster cooperative learning among students, where they are eager to help and evaluate their peers and take ownership of their own language learning success. This can lead to improvement of social skills, better evaluation, and more accurate evaluations. The social interdependence theory, which holds that students help one another learn better because they care about the group and its members and want to accomplish the same goal, supports our findings (Slavin, 2011). According to the social interdependence hypothesis, the students' collaboration can aid them in achieving their common goals. Our research also supports Vygotsky's social constructivism theory, which contends that cooperative learning activities are beneficial for students because they allow 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).

Receiving additional input and feedback is another factor that contributed to the study's improved outcomes. The peer-evaluation group students received more input and comments, which improved their speaking and writing abilities. Peer evaluation also gave the students a greater sense of agency in their language acquisition. Peer

assessment helped students become more conscious of classroom discussions since they were required to evaluate both their own performances and those of their classmates. As a result, they were more attentive to the assessment procedure as a method of learning and language development.

The success of the self-assessment group on the writing and speaking posttests may have resulted from the use of self-assessment, which encouraged students to learn independently to meet their learning objectives and enhance their capacity for future performance. According to Fraenkel et al. (2011), self-assessment encourages students to improve and succeed in meeting all requirements by enabling them to see how much they have achieved the learning goal. Self-evaluation may also encourage pupils to take on more responsibility for enhancing their speaking performance. This is important as it is endorsed by Ma and Winke (2019) who stated that self-assessment allows the learners to have some consciousness about their duty towards the learning objectives in term of students' speaking skills.

Briefly stated, the results of this study demonstrate that in contrast to common methods of teaching writing and speaking skills, which may not give students the chance to evaluate their own or their peers' performance, both self- and peer assessment can significantly improve speaking and writing skills of EFL learners. Therefore, it is strongly advised that language teachers, especially those who focus on speaking and writing skills, incorporate more educational practices like self- and peer assessment into their lessons.

Many different elements might contribute to learners' language learning. Assessment might be one of them. With both self- and peer assessment, the students actively participate in the evaluation procedure and assist to foster language learning. Self- and peer assessment encourage lifelong learning by assisting students in objectively assessing their own and their peers' accomplishments. Peer assessment helps students feel like they are a part of a community of academics since it allows them to engage in an important component of higher education while also providing a critical evaluation of the work of others.

This research can have some implications for EFL teachers, learners, and syllabus designers. Language learners can become accustomed to many forms of examinations, including those that were the subject of the current study, with the help of EFL teachers and instructors. Teachers can improve their assistance of students by using the evaluative feedback obtained through peer and self-assessments. Peer and self-assessment are useful tools for the development of speaking and writing skills, and teachers may find it advantageous to use them. In addition to using a variety of assessments in the classroom, teachers can encourage students' autonomy by having them participate in self-assessment activities. As a result, EFL students can gradually grasp what it takes to become self-directed learners. Peer-evaluation activities are another way that teachers can promote cooperative learning.

As for EFL students, the outcomes might recommend that being familiar with diverse kinds of assessment in general and self-assessment and peer assessment in particular would assist them to develop their speaking and writing skills in the same manner. Also, by using self-assessment activities, students can recognize the exact areas in which they need assistance and support, and then, they can ask help from

their teachers. Furthermore, since the positive effects of self-assessment and peer-assessment on EFL learners speaking and writing skills were found in this research, teaching methods and material development in language classes can be designed in such a way that encourage self-assessment techniques and accordingly influence foreign language learners' language skills.

Learners typically have more to gain from self- and peer-assessment settings than they do from having their work graded by a tutor. They gain knowledge via doing assessments and frequently from receiving oral as well as written feedback. To ensure that all students receive fair and equal treatment, the tutor should monitor the feedback and, where necessary, elaborate it. The assessment criteria can be created by the tutor, but it adds more value when the students are actively participating in the process. The students participate in the group during peer assessment and exhibit leadership skills.

Material developers and syllabus designers are also the beneficiaries of the current investigation. The results can assist syllabus designers to grasp a better picture of three assessments and how they can affect EFL learners' language learning. As peer- and self-assessment sorts were found to be useful in the same way on EFL learners' speaking and writing skills, syllabus designers are suggested to integrate various assessment kinds in their syllabi. The outcomes of this research may also help material developers in preparing various activities and tasks adequately suitable to the EFL students' speaking and writing skills.

Like other research, this one had some limitations and was unable to cover all the pertinent problems.

1. The study's inclusion of only individuals between the ages of 17 and 25 is one of its limitations. As a result, the findings cannot be applied to other age groups.
2. There were no more than 75 persons in the present study. This cannot, therefore, be applied generally.
3. Because only female students were included in the study, it is possible that the findings do not apply to male students.

There are a few recommendations made for additional research. It is suggested that the treatment of the present study be repeated over a prolonged length of time in the next studies. To give the researcher(s) more precise and broadly applicable results, the same study can be conducted on a bigger sample of EFL students. It is possible to do research on the impact of peer and self-evaluation on other language proficiency. As this research did not include a delayed posttest, future studies can fill this gap by examining the long-term effects of self-, teacher, and peer- assessment on language learning. Additionally, qualitative research on the use of various assessment types and their effects on students and teachers can be done by conducting interviews and observing what happens in the classroom.

Authors' information
Anwar Hammad Al-Rashidi got her PhD in Counseling Psychology from Imam Mohamed bin Saud Islamic University. She is an Associate Professor at the Department of Psychology in Prince Sattam Bin Abdulaziz University, Al-Kharj, Saudi Arabia.
Muhammad Asif is an Assistant Professor at the Department of English and Literary Studies, University of Management and Technology, Lahore, Pakistan. His research interests include learning technologies, language testing and assessment, second language acquisition, and web-delivered language learning platforms.
Masoumeh Ghasemi Vanani held a master degree of language teaching from Islamic Azadi University of Najafabad. She graduated in 2009. She has been working as English teacher at ILI since 2011. According to her years of experience as English teacher, she felt two important productive skills of writing and speaking have not received enough attention, and they are the most difficult skills in students' view. Since writing skill is important for education, immigration, and finding job opportunities, therefore, she likes to find ways to improve these two productive skills.
Amhara Aberash got her PhD in TEFL from Jimma University of Ethiopia. She is now an independent researcher and published many papers in different journals.

Availability of data and materials
The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Competing interests
The authors declare that they have no competing interests.

References

Abdul, G. M., Nur, A., & Ismail, S. W. (2018). Improving the reading comprehension through cognitive reading strategies in language class of coastal area in Indonesia. IOP Conf. Series: Earth and Environmental Science, *156*, 1–10. https://doi.org/10.1088/1755-1315/156/1/012050.

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.

Ahıskalı, E. (2020). Planlama ve Yazma Stratejileri Öğretiminin Planlama ve Yazma Başarısına, Yazma Özyeterliğine ve Yazma Kaygısına Etkisi. Dokuz Eylül Üniversitesi.

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.

Al-Mahrooqi, R., & Denman, C.J. (2018). Alternative forms of assessment. In J.I. Liontas (M. DelliCarpini, C. Coombe), *The TESOL encyclopedia of English language teaching* (1st ed.), vol. 8. Hoboken: Wiley. 4851–4856.

Ariafar, M., & Fatemipour, H. R. (2013). The effect of self-assessment on Iranian EFL learners' speaking skill. *International Journal of Applied Linguistics and English Literature*, *2*(4), 7–15.

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 & Practice in Language Studies*, *5*(5), 1–18.

Birjandi, P., & Siyyari, M. (2010). Self-assessment and peer-assessment: A comparative study of their effect on writing performance and rating accuracy. *Language Testing*, *3*(2), 10–23.

Bora, F. D. (2012). The impact of emotional intelligence on developing speaking skills: From brain-based perspective. *Procedia-Social and Behavioral Sciences*, *46*, 2094–2098.

Bourke, R., & Mentis, M. (2011). Self-assessment as a process for inclusion. *International Journal of Inclusive Education*, *17*(8), 854–867.

Brown, J. D., & Hudson, T. (2012). The alternatives in language assessment. *TESOL Quarterly*, *32*(4), 653–675.

Brown, S., & Glasner, A. (2007). *Evaluar en la Universidad. Problemas y nuevos enfoques*. Madrid: Narcea S. A. Ediciones.

Butler, Y. G., & Lee, J. (2010). The effects of self-assessment among young learners of English. *Language Testing*, *27*(1), 5–31.

Celce-Murcia, M. (2001). *Teaching English as a second or foreign language*, (3rd ed.,). Boston: Heinle & Heinle.

Chen, Y.S. (2008) The Positive Effect of Green Intellectual Capital on Competitive Advantages of Firms. *Journal of Business Ethics, 77,* 271–286. https://doi.org/10.1007/s10551-006-9349-1.

Coombe, C., Folse, K., & Hubley, N. (2007). *A practical guide to assessing English language learners*, (1st ed.,). Ann Arbor: University of Michigan.

Dhindsa, H., Omar, K., & Waldrip, B. (2007). Upper secondary Bruneian science students' perceptions of assessment. *International Journal of Science Education*, *29*(10), 1261–1280.

Dikel, M. R. (2009). *A guide to going online for self-assessment tools* https://www.vault.com/blogs/job-search/aguide-to-going-online-for-self-assessment-tools.

Ellis, R. (2004). Tasks, production & lagunage acquisition. In R. Ellis (Ed.), *Task-based language learning and teaching*, (pp. 10–140). Oxford: Oxford University Press.

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2011). *How to design and evaluate research in education*. New York: McGraw-Hill Humanities/Social Sciences/Languages.

Fulcher, G., & Davidson, F. (2007). *Language testing and assessment: An advanced resource book*. London & New York: Routledge.

Harris, K. R., Graham, S., & Adkins, M. (2015). Practice-based professional development and self-regulated strategy development for tier 2, at-risk writers in second grade. *Contemporary Educational Psychology*, *40*, 5–16.

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.

Harris, M., & McCann, P. (1994). *Assessment*. London: Heinemann.

Heidarian, N. (2016). Investigating the effect of using self-assessment on Iranian EFL learners' writing. *Journal of Education and Practice*, *7*(28), 80–89.

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.

Hidi, S., & Boscolo, P. (2006). Motivation and writing. In C. A. MacArthur et al. (Eds.), *Handbook of writing research*. The Guilford Press.

Hughes, A. (2003). *Testing for language teachers*, (2nd ed.,). Cambridge: Cambridge University Press.

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.

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.

Jafarigohar, M. (2017). The effect of assessment technique on EFL learners' writing motivation and self-regulation. *Iranian Journal of English for Academic Purposes*, *9*(4), 141–162.

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.

Kara, A. (2009). The effects of a learning theories unit on students' attitudes towards learning. *Australian Journal of Teacher Education*, *34*(3), 100–113.

Karatay, H. (2013). In M. Özbay (Ed.), *Süreç Temelli Yazma Modelleri: Planlı Yazma ve Değerlendirme. Yazma Eğitimiiçinde*, (pp. 21–45). Pegem Akademi.

Keith, J. (2005). Trends in peer learning. *Educational Psychology*, *25*(6), 631–645.

Kurudayıoğlu, M., & Özay, K. (2010). Examining the written expressions of primary school graders as regard to their choice of topics. *Mustafa Kemal University Journal of Social Sciences Institute*, *7*(13), 192–207.

Linn, R. L., & Miller, M. D. (2005). *Measurement and assessment in teaching (9thed.)*. Upper Saddle River: Prentice Hall.

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.

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. https://doi.org/10.15548/jt.v27i2.613.

Matsuno, S. (2009). Self-, peer, and teacher assessments in Japanese university EFL writing classrooms. *Language Testing*, *26*(1), 75–100.

Mazloomi, S. & Khabiri, M. (2018). The impact of self-assessment on language learners' writing skill. *Innovations in Education and Teaching International, 55*(1), 91-100. https://doi.org/10.1080/14703297.2016.1214078.

Montano, D. E., & Kasprzyk, D. (2008). Theory of reasoned action, theory of planned behavior, and the integrated behavioral model. In *Theory, research, and practice*. San Francisco: Jossey-Bass.

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.

Nunan, D. (2003). *Practical English language teaching*. New York: McGraw-Hill.

Ounis, A. (2017). The assessment of speaking skills at the tertiary level. *International Journal of English Linguistics*, *7*(4), 95–109.

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

Ramani, S., Konings, K. D., Mann, K. V., Pisarski, E. E., & van der Vleuten, C. P. M. (2018). About politeness, face, and feedback: Exploring resident and faculty perceptions of how institutional feedback culture influences feedback practices. *Academic Medicine*, *93*(9), 1348–1358.

Rashidi, N., & Nazemi, H. (2015). The effect of Iranian EFL learners' attitude towards English language learning on their autonomy. *Journal of Studies in Learning and Teaching English*, *4*(2), 133–148.

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., & Renandya, W. A. (Eds.) (2002). *Methodology in language teaching: An anthology of current practice*. Cambridge: Cambridge university press.

Richards, J. C., & Schmidt, R. (2002). *Dictionary of language teaching and applied linguistics*. London: Pearson Education Limited.

Richards, J. C., & Schmidt, R. (2010). *Longman dictionary of language teaching and applied linguistics*, (4th ed.,). London: Pearson Longman.

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.

Schut, S., Driessen, E., van Tartwijk, J., van der Vleuten, C., & Heeneman, S. (2018). Stakes in the eye of the beholder: An international study of learners' perceptions within programmatic assessment. *Medical Education*, *52*(6), 654–663.

Shafiee Rad, H., Namaziandost, E., & Razmi, M. H. (2022). Integrating STAD and flipped learning in expository writing skills: Impacts on students' achievement and perceptions. *Journal of Research on Technology in Education*, *1–17*. https://doi.org/10.1080/15391523.2022.2030265.

Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, *29*(7), 77–84.

Slavin, R. E. (1997). *Educational psychology: Theory and practice*. Allyn & Bacon: Needham Heights.

Slavin, R. E. (2011). Instruction based on cooperative learning. In R. E. Mayer, & P. A. Alexander (Eds.), *Handbook of research on learning and instruction*, (pp. 344–360). New York: Taylor & Francis.

Spratt, M., Pulverness, A., & Williams, M. (2008). *The tkt. Teaching knowledge test course*. Cambridge: University of Cambridge. ESOL examinations.

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.

Telio, S., Ajjawi, R., & Regehr, G. (2015). The educational alliance as a framework for reconceptualizing feedback in medical education. *Journal of the Association of American Medical Colleges*, *90*(5), 609–614.

Tigchelaar, M. (2019). Exploring the relationship between self-assessments and OPIc ratings of oral proficiency in French. In *Dalam foreign language proficiency in higher education (hlm. 153–173)*. Springer.

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

Topuzkanamış, E. (2014). *Yazma Stratejileri Öğretiminin Türkçe Öğretmenliği Birinci Sınıf Öğrencilerinin Yazılı Anlatım Başarısı ve Yazma Kaygısına Etkisi*. Gazi University.

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.

Ünaldı, İ. (2016). Self and teacher assessment as predictors of proficiency levels of Turkish EFL learners. *Assessment & Evaluation in Higher Education*, *41*(1), 67–80.

Watling, C., & LaDonna, K. (2019). Where philosophy meets culture: Exploring how coaches conceptualize their roles. *Medical Education*, *53*(5), 467–476.

Weaver, D., & Esposto, A. (2011). Peer assessment as a method of improving student engagement. *Assessment and Evaluation in Higher Education*, *37*(7), 805–816.

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

Wiggins, G., & McTighe, J. (2005). *Understanding by design: Expanded (2nd ed.)*. Alexandria: ASCD.

Wigglesworth, G., & Storch, N. (2009). Pair versus individual writing: Effects on fluency, complexity, and accuracy. *Language Testing*, *26*(3), 445–466. https://doi.org/10.1177/0265532209104670.

Wood, W. B. (2009). Innovations in teaching undergraduate biology and why we need them. *Annual Review of Cell and Developmental Biology*, *25*, 93–112.

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.

Zare Toofan, Z., Vaseghi, R., & Zare, M. (2019). Iranian EFL learners' perceptions toward paper assessment in mid-term and final exams in an English language institute. *International Journal of Research in English Education (IJREE)*, *4*(3), 21–41.

Publisher's Note