

Paper Type: Original Article



Exploring Teachers' Practices in Designing Writing and Speaking English Language Tests at Iranian Junior High Schools

Firooze Foroozan^{1,*}, Mohammad Javad Rezai²

¹ Department of English Language and Literature, Faculty of Languages and Literature, Yazd University, Yazd, Iran;

² Department of English Language and Literature, Faculty of Languages and Literature, Yazd University, Yazd, Iran.

Received: 18 November, 2025

Revised: 21 March, 2026

Accepted: 18 May, 2026

Abstract

Despite curriculum reforms aimed at fostering communicative competence in English education, the implementation of communicative assessment practices remains a challenge in Iranian junior high schools. To explore this issue, this study investigated how EFL teachers design writing and speaking assessments in line with the goals of the Prospect series, introduced in 2013 as part of Iran's national English curriculum. Addressing a critical gap in the existing research, the study examined EFL teachers' practices and challenges in aligning their tests with communicative language teaching (CLT) principles using an explanatory sequential mixed methods design. In the quantitative phase, the data were collected from 100 junior high school English teachers through a structured questionnaire. In the qualitative phase, 10 experienced EFL teachers were chosen through purposive sampling for in-depth semi-structured interviews. The findings revealed a disconnection between curriculum goals and classroom assessment practices. The key challenges included large class sizes, insufficient training in CLT-oriented assessment design, limited access to teaching and testing resources, and time constraints. The teachers also expressed a strong need for focused professional development tailored to writing and speaking test design. The study offers implications for policymakers and curriculum developers; they should provide practical assessment models and rubrics aligned with communicative objectives.

Keywords: Assessment challenges, Communicative language testing, Curriculum reform, Iranian education system, Productive language skills.

I | INTRODUCTION

Assessment plays a fundamental role in English as a Foreign Language (EFL) education, serving as a critical tool for evaluating students' progress, identifying learning gaps, and informing instructional practices (Popham, 2011). In language learning, assessments are not merely instruments for measuring knowledge. They shape pedagogical decisions and influence students' motivation and achievement (Fulcher, 2012). Within this context, writing and speaking assessments hold particular importance as they directly impact students' ability to communicate effectively in English. However, the design and implementation of these assessments remain challenging for many teachers, particularly in centralized educational systems such as Iran, where curriculum changes and policy constraints shape assessment practices (Atai & Mazlum, 2013; Razavipour & Rezagah, 2018).



Licensee

Journal of Studies in Language Learning and Teaching. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY-NC) license.



Corresponding Author: firoozeforoozan@gmail.com

Cite this article: Foroozan, F., Rezai, M. J., (2026). Exploring Teachers' Practices in Designing Writing and Speaking English Language Tests at Iranian Junior High Schools, *Journal of Studies in Language Learning and Teaching*, 2(2), 105-119. doi: 10.22034/JSLLT.2026.23994.1103

In the Iranian junior high school system, English education and major aspects like school systems, curriculum standards, textbooks, and exams are managed by the Ministry of Education (Atai & Mazlum, 2013). Historically, English language testing in Iran has primarily emphasized grammar, reading comprehension, and vocabulary development, with limited attention given to productive skills such as speaking and writing (Sadeghi & Richards, 2016). The Grammar-Translation Method (GTM) dominated the instructional approach, leading to generations of students with high grammatical competence but limited communicative abilities (Zayniddinova, 2025). Consequently, students were not tested on speaking and listening. According to Sadeghi and Richards (2016), most Iranian students who completed six years of English instruction in high school demonstrated limited proficiency in listening and speaking.

In response to these limitations, the Ministry of Education introduced a series of curriculum reforms, including the "Prospect" textbooks, to promote communicative competence by incorporating listening and speaking activities into instruction and assessment (Khadrir-Sharabyan et al., 2014). These reforms mandated that speaking and listening skills be included in the final exam, contributing to a portion of the overall score. However, educational reforms, particularly in language education, are complex processes that involve multiple stakeholders, each bearing distinct expectations, experiences, and constraints.

Teachers are the primary agents in designing and administering language tests, yet their assessment practices are shaped by several factors, including pedagogical beliefs, professional training, and classroom realities (Scarino, 2013). Research suggests that many Iranian EFL teachers lack adequate training in designing communicative assessments, and they often rely on traditional testing methods that are not fully in line with the curriculum objectives (Kardoust & Saeedian, 2021). Speaking tests, for instance, may be reduced to memorized dialogues or isolated pronunciation drills rather than holistic assessments of communicative ability. Writing tests, similarly, often prioritize grammatical accuracy over coherence, organization, and content development. These discrepancies highlight a misalignment between assessment practices and the communicative goals of the new curriculum, underscoring the need for targeted professional development in assessment literacy (Vogt & Tsagari, 2014).

Another critical challenge in test design is the influence of external constraints such as large class sizes and time limitations (Nguyen & Le, 2021). Iranian junior high school teachers might face logistical difficulties in administering individual speaking assessments, leading them to adopt alternative evaluation methods that may not accurately reflect students' communicative competence (Noughabi, 2017). Writing assessments, too, are constrained by grading workloads, prompting some teachers to favor short-answer responses over more comprehensive writing tasks. These practical concerns, combined with the pressure of standardized testing, contribute to a washback effect in which teaching to the test takes precedence over authentic language use (Riazi & Razavipour, 2011).

Despite these challenges, effective assessment of writing and speaking skills is essential for fostering students' communicative competence. Well-designed assessments can enhance students' motivation to learn, as they provide clear goals and immediate, constructive feedback on their progress (Mellati & Khademi, 2018). Therefore, the present study aimed to investigate the assessment practices of Iranian junior high school English teachers in designing writing and speaking tests. By examining teachers' perceptions, strategies, and challenges, this research sought to bridge the gap between curriculum objectives and classroom realities. Understanding how teachers interpret and implement assessments can offer valuable insights into the effectiveness of current testing practices.

II. REVIEW OF THE RELATED LITERATURE

2.1. Language assessment

Language assessment has long been a critical component of language education. It functions both as a tool for evaluating learners' language proficiency and as a means of guiding instruction. The distinction between testing and assessment is essential to understanding this field. While tests are formal evaluations that occur at specific times, assessment is an ongoing process that occurs every time a learner interacts

with language in the classroom (Brown & Abeywickrama, 2010, p. 3). This continuous nature of assessment is crucial for providing feedback that can shape both teaching and learning.

Language assessment has historically been an essential element of language education, functioning as both a method for evaluating learners' linguistic proficiency and a guide for instructional strategies. The differentiation between testing and assessment is crucial for comprehending this domain. Tests are formal evaluations conducted at designated intervals, whereas assessment is a continuous process that takes place whenever a student engages with language in the classroom (Brown & Abeywickrama, 2010). The ongoing nature of assessment is essential for delivering feedback that can influence both instruction and learning. In recent years, the concept of assessment literacy has become central to understanding effective teaching practices, particularly in contexts of curriculum reform. Assessment literacy encompasses not only technical knowledge of test design, validity, and reliability, but also an understanding of the social, ethical, and pedagogical implications of assessment (Inbar-Lourie, 2016; Taylor, 2013). As Stiggins (1991) argued early on, teachers must be able to generate and use quality assessments to improve instruction and support student learning. In language education, this means being equipped to design tasks that reflect communicative competence and real-world language use, rather than relying solely on discrete-point testing. Scholars emphasize that teachers' assessment decisions are deeply influenced by their beliefs, training, and contextual constraints, all of which are facets of their assessment literacy (Fulcher, 2012; Scarino, 2013).

Recent studies have explored the impact of assessment literacy on teaching and learning outcomes. In a study by Xu and Brown (2016), it was found that teachers with higher levels of assessment literacy were better able to align their assessments with instructional goals and provide more targeted feedback to students. This alignment, in turn, contributed to improved student learning outcomes and greater student engagement with the assessment process. In the Iranian context, limited pre-service and in-service training opportunities frequently result in gaps in assessment literacy. Consequently, EFL teachers face difficulties in designing assessments that reflect communicative curriculum goals (Aliakbari, Yasini, & Sadeghi, 2023). Therefore, improving assessment literacy is essential for achieving meaningful alignment between classroom assessment and pedagogical objectives.

Authentic assessment refers to tasks that closely mirror real-world language use, requiring learners to apply their language skills in meaningful contexts (Gulikers et al., 2004). This approach contrasts with traditional language tests, which often focus on discrete language skills in isolation from communicative use. Proponents of authentic assessment argue that it provides a more accurate representation of a learner's language abilities and better prepares them for real-life communication challenges (Lantolf & Poehner, 2014).

2.2. Communicative language testing

Communicative language testing has become an essential method in language evaluation, closely associated with the principles of communicative language instruction (Kaçar, 2023). Unlike traditional forms of language testing that prioritize the assessment of discrete grammatical or lexical knowledge, it is designed to evaluate a test-taker's ability to use language effectively in real-life communicative situations. This paradigm shift reflects broader changes in language education, where the focus has increasingly moved towards practical language use, interaction, and the capacity to navigate various social contexts (Yuzar, 2020). Communicative language testing emerged in response to the limitations of traditional testing methods that predominantly used multiple-choice items, focusing on reliability at the expense of validity. In the 1980s, language testers began to question whether these tests could adequately assess a learner's ability to communicate in the target language. The resulting communicative testing movement emphasized tasks that mirrored real-life language use, such as role plays, interviews, and problem-solving activities (Fulcher, 2000).

Recent research continues to emphasize the importance of communicative competence in language assessment. According to Fulcher (2015), the evolution of communicative language testing reflects a growing recognition of the need to assess language in a way that mirrors its use in real-world contexts. This involves not only evaluating linguistic accuracy but also the ability to perform communicative functions such as requesting information, apologizing, or negotiating meaning in interaction. The shift towards performance-based assessments, where test-takers engage in tasks that simulate authentic communicative situations, is a hallmark of communicative language testing and distinguishes it from more traditional testing methods.

Sa'diyah (2020) highlights that English teachers often employ alternative assessment methods, such as peer evaluations and role-playing, to assess speaking skills.

Recent research in Iran, such as the study by Tayyebi, Abbasabady, and Abbassian (2022), indicates that although Iranian EFL teachers are familiar with the principles of writing task design in theory, they often rely on traditional assessments and have limited training in designing authentic writing tasks. In terms of speaking assessment, Namaziandost, Hashemifardnia, and Shafiee (2019) explored how task-based activities, such as simulated real-life situations (e.g., ordering food, visiting a doctor), can enhance speaking fluency. However, their experimental design did not examine the specific tasks that teachers use in routine classroom assessments, which leaves a gap in understanding how authentic teacher-designed tasks affect student performance. Similarly, Karizak, Afraz and Samimi (2024) developed a self-assessment tool for speaking but did not investigate how teachers design speaking tasks in classroom assessments. Furthermore, Nevisi and Hosseinpour (2022) investigated task-based speaking assessments in Iranian academic settings, revealing the use of both summative and formative speaking tasks. However, their focus was limited to university contexts and not junior high schools, which further illustrates the absence of studies targeting secondary education in Iran regarding writing and speaking task design.

2.3. Language assessment in secondary education in Iran

In recent years, there has been a noticeable shift in language assessment practices in Iranian secondary education, moving from traditional methods towards more communicative and performance-based assessments (Aliakbari et al., 2023). This change is in line with global trends that emphasize the importance of assessing students' ability to use language in real-world contexts rather than simply testing their knowledge of grammatical rules and vocabulary. A study by Razavipour and Rezagah (2018) highlights that Iranian educational policymakers have increasingly recognized the limitations of traditional language assessments, which often rely heavily on multiple-choice questions and rote memorization. In response, there has been a push towards incorporating tasks that require students to demonstrate their speaking, listening, writing, and reading skills in more authentic and meaningful ways. Despite these efforts, the implementation of communicative language assessment in Iranian secondary schools has been inconsistent. One of the primary challenges is the lack of teacher training in new assessment methods. As documented by Farhady and Tavassoli (2021), many teachers face overcrowded classrooms, a lack of assessment materials, and escalating administrative responsibilities. This is further confirmed by Aliakbari et al. (2023), who found that Iranian EFL teachers often experience a significant discrepancy between theoretical knowledge of communicative assessment and its application in actual classroom settings. This has led to a situation where teachers continue to rely on traditional testing methods, even when the curriculum emphasizes communicative competence. The study also points out that the high-stakes nature of national exams reinforces this reliance on traditional assessments, as teachers and students alike are focused on preparing for these exams, which are predominantly multiple-choice and do not adequately assess communicative skills.

The current study sought to answer the following research questions based on the above-mentioned research background.

1. How do Iranian junior high school EFL teachers design writing tests to assess communicative competence in line with the new curriculum's requirements?
2. How do Iranian junior high school EFL teachers design speaking tests to assess communicative competence in line with the new curriculum's requirements?
3. What challenges do Iranian EFL teachers face when designing writing and speaking tests?

III. METHODOLOGY

3.1. Research design

The current study was performed in two phases following an explanatory sequential mixed methods design (Creswell & Creswell, 2017). This design was chosen to gain both broad, quantifiable insights and deep, contextual understanding of the teachers' perspectives. The quantitative phase involved administering a questionnaire, which provided statistical data about the teachers' test design practices, while the qualitative phase involved semi-structured interviews that validated the quantitative data and allowed for in-depth exploration of teachers' experiences and the challenges they face. This combination

of methods aimed to provide a comprehensive view of the research problem, leveraging the strengths of both approaches (Bryman, 2006; Creswell, 2013).

3.2. Participants

In the quantitative phase, a total of 150 junior high school English teachers initially participated in the study on a voluntary basis. However, 50 participants were excluded due to having less than five years of teaching experience, resulting in a final sample of 100 Iranian junior high school EFL teachers. These teachers were selected using convenience sampling, a common non-probability technique, based on their accessibility and willingness to participate (Stratton, 2021). The participants were aged between 25 and 54 years ($M = 38.01$, $SD = 8.03$) and had varying educational qualifications, including Bachelor's (34%), Master's (56%), and Doctoral (10%) degrees. Gender distribution was relatively balanced, with 44% male and 56% female participants. The teachers' professional experience ranged from 6 to over 15 years ($M = 14.12$, $SD = 6.75$), with the majority having 6-10 years of experience (52%). They were from several regions across Iran, including Yazd (48%) and Tehran (26%), as well as Shiraz (8%), Mashhad (5%), Kerman (4%), Gorgan (3%), Isfahan (3%), Zahedan (2%), and Qazvin (1%), representing a diverse sample of EFL teachers.

For the qualitative phase, purposive sampling was used to select 10 teachers, six females and four males. The participants were drawn from a larger pool of 100 teachers and were selected because they had more than 20 years of experience and were actively involved in designing high-stakes assessment items. All had served as test designers for provincial-level examinations, making their background directly relevant to the study's focus on writing and speaking assessment.

3.3. Instruments

The study employed two main instruments: A questionnaire and semi-structured interviews.

3.3.1 Questionnaire on designing writing and speaking skills

The questionnaire, developed by the researchers of the current study, consisted of 26 Likert-scale items targeting teachers' practices in designing writing and speaking tests. A separate section collected demographic information such as age, gender, educational qualification, city, and teaching experience. To ensure content validity, the questionnaire was reviewed by three experts in language assessment and test design. All the experts verified the content of the questionnaire along with a number of suggestions for improvement. Additionally, a pilot study was conducted with 12 experienced English teachers holding BA and MA degrees, each with over 15 years of teaching experience. Based on the feedback from both the expert reviewers and pilot participants, several items were revised, some were deleted, and a few items were edited for greater clarity and relevance.

To establish construct validity, the items of the questionnaire were subjected to principal components analysis (PCA) using *IBM SPSS Statistics* (Version 26). Prior to performing PCA, the suitability of the data for factor analysis was examined. Inspection of the correlation matrix revealed the presence of many coefficients of .30 and above. The Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy was .893, exceeding the recommended threshold of .60 (Kaiser, 1970, 1974), and Bartlett's Test of Sphericity (Bartlett, 1954) reached statistical significance, indicating that the correlation matrix was factorable. PCA initially revealed the presence of four components with eigenvalues greater than 1, explaining 44.8%, 6.8%, 6.7%, and 4.8% of the total variance, respectively. However, visual inspection of the scree plot revealed a clear break after the first component. Following Cattell's (1966) scree test, only one component was retained for further analysis. This decision was further supported by the results of parallel analysis, which indicated that only one component had an eigenvalue exceeding the corresponding criterion value from a randomly generated data matrix of the same size (23 variables $\times$ 100 respondents). These results suggest that the questionnaire items primarily loaded onto a single underlying construct, i.e. productive language ability. This, in turn, supports the unidimensionality of the instrument and confirms its construct validity in measuring teachers' practices in designing communicative writing and speaking assessments.

The reliability of the revised questionnaire was then evaluated using Cronbach's alpha coefficient. The overall reliability was found to be excellent, with a coefficient of $\alpha = .943$ (Hinton, McMurray, & Brownlow, 2014). Detailed reliability values for each section, including the number of items and corresponding Cronbach's alpha values, are presented in Table 1 below.

Table 1. Questionnaire details and reliability

Number of items	Name of section	Cronbach's alpha
11	Writing skill assessment	0.874
15	Speaking skill assessment	0.904

3.3.2 Interview

To complement the quantitative data, semi-structured interviews were conducted with a subset of the teachers ($N = 10$) selected through purposive sampling. The interviews aimed to explore teachers' experiences and challenges in designing writing and speaking tests in more detail. Sample questions included:

1. What challenges do you face when designing writing tests?
2. What challenges do you face when designing speaking tests?
3. How much training have you received in designing effective language assessments?
4. What additional support or resources would be most helpful for you to improve your test design skills?

Each interview lasted 25-30 minutes and was conducted via Google Meet. The interviews were recorded, transcribed, and analyzed thematically to identify key themes and patterns.

3.4. Data collection procedure

Data collection was carried out in two phases. In the first phase, the questionnaire was distributed to 100 teachers, who completed it either online via Google Forms or on paper. After gathering the responses, the quantitative data were analyzed using SPSS (version 26). In the second phase, semi-structured interviews were conducted with 10 teachers who had already completed the questionnaire. The interviews provided deeper insights into the teachers' experiences and practices. Each interview was transcribed verbatim, and the data were analyzed using inductive thematic analysis to identify recurring themes and patterns related to test design.

Before data collection began, all participants were informed about the purpose of the study and their rights as participants. Informed consent was obtained from both questionnaire respondents and interview participants, who were assured that their participation was entirely voluntary and that they could withdraw at any time without penalty. Participants' anonymity and confidentiality were strictly maintained and no identifying information was collected or disclosed. The data were stored securely and used solely for research purposes.

3.5. Data analysis

The quantitative data from the questionnaire were analyzed using SPSS version 26. Descriptive statistics, such as means and frequencies, were calculated for each item to summarize teachers' responses. A chi-square goodness of fit test was conducted to find out whether the respondents had a similar preference for the different options. The qualitative data from the interviews were analyzed using thematic analysis (Braun & Clarke, 2006). The interview transcriptions were coded, and recurring themes related to writing and speaking test design were identified. The coding was confirmed via intra-rater reliability, which indicated complete agreement ($r = 1.00$). The analysis was conducted iteratively, ensuring that the themes were consistently relevant to the research questions.

IV. RESULTS

4.1. Analysis of the teacher questionnaire

The first phase of the study was quantitative in nature. A questionnaire was developed to gather data from the teachers, focusing on their practices and perspectives in designing writing and speaking English language tests. The data collected from the 100 Iranian EFL teachers who responded to 26 items in the researcher-made questionnaire were fed into SPSS.

4.1.1. Writing skill

The first section of the questionnaire aimed to understand the teachers' practices in designing writing tests. Among the various assessment formats, short-answer tasks emerged as the most frequently used method, with 79% of the teachers incorporating them in their assessments. Paragraph writing was also commonly used, with 58% of the teachers selecting this method. Sentence completion tasks were utilized

by 57% of the teachers, while picture-cued tasks and ordering tasks were moderately used by 45% and 40% of the teachers, respectively. The least common method was Dicto-Comp (Dictation-Composition), used by only 30% of the teachers. Figure 1 shows the percentage of each assessment format selected by the teachers.

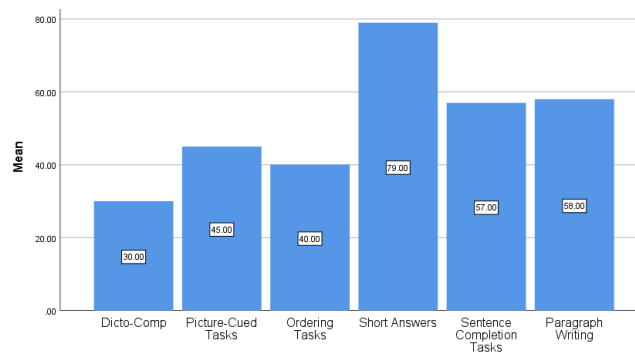


Figure 1. Percentages of the assessment formats used by the teachers to test writing

The teachers' views on writing test design were also assessed in the questionnaire. The responses indicated a general alignment with curriculum goals and the importance of clear instructions. The item with the highest mean score (4.08) indicates that teachers often provide clear instructions for writing tasks. However, many teachers reported only sometimes (43%) or rarely (16%) designing writing tests that focus on communicative skills, as reflected in the lower mean score of 3.33 for the first item. The survey also revealed that the teachers generally assessed language accuracy in students' writing, with 74% doing it so frequently. Furthermore, the teachers reported feeling confident in their ability to design effective writing tests ($M = 3.75$), with 65% indicating that they felt confident most or all of the time. Table 2 shows the teachers' views on their assessment practices for writing.

Table 2. Teachers' views on designing writing tests

Context	Mean	Never (%)	Rarely (%)	Sometimes (%)	Often (%)	Always (%)	Chi-Square	df	Sig.
1. I design writing tests that focus on communicative skills.	3.33	1	16	43	29	11	53.40	4	0.000
2. I make sure my writing tests cover the learning goals from the curriculum.	3.99	0	2	27	41	30	32.56	3	0.000
3. My writing tests are related to students' interests and experiences.	3.63	0	8	35	43	14	33.36	3	0.000
4. I provide clear instructions for each writing task.	4.08	2	3	18	39	38	56.10	4	0.000
5. I use already-prepared tests from different sources to assess students' writing ability.	3.46	3	6	48	28	15	67.90	4	0.000
6. I ensure the suitability of prepared tests before using them.	4.04	1	3	22	39	35	62.00	4	0.000
7. I assess language accuracy in students' writing.	3.91	2	4	20	49	25	72.30	4	0.000
8. I provide feedback on students' writing to help them improve.	3.88	1	5	25	43	26	58.80	4	0.000
9. I regularly assess students' progress in writing and adjust my teaching accordingly.	3.65	2	2	45	31	20	69.70	4	0.000
10. I feel confident in my ability to design effective writing tests.	3.75	2	5	28	46	19	64.50	4	0.000

4.1.2. Speaking skill

The analysis of speaking assessment practices revealed that the most commonly used formats for evaluating speaking were question-and-answer formats (77%) and role-plays (74%). These methods were followed by interviews (59%), reading aloud (50%), and presentations (47%). The least common method was monologues, used by only 2% of the teachers. Figure 2 shows the percentage of each assessment format implemented by the teachers.

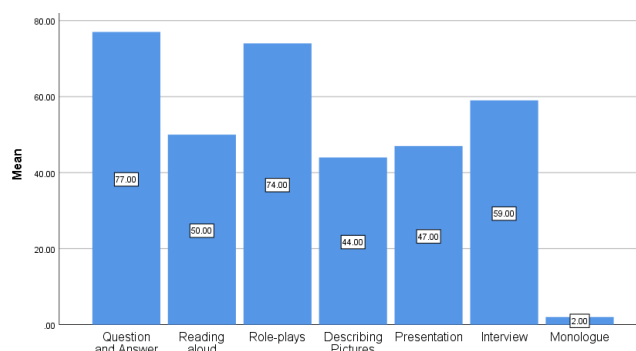


Figure 2. Percentages of assessment formats used by the teachers to test speaking

According to the responses regarding assessment methods, classroom observation and activities were the most favored method to evaluate speaking skills used by (77%) of the teachers. Following closely, speaking tests were used by (69%) of the teachers. These structured tests effectively assess various aspects of speaking, including pronunciation, fluency, and the ability to respond to questions and prompts. Figure 3 shows the percentage of each assessment method selected by the teachers.

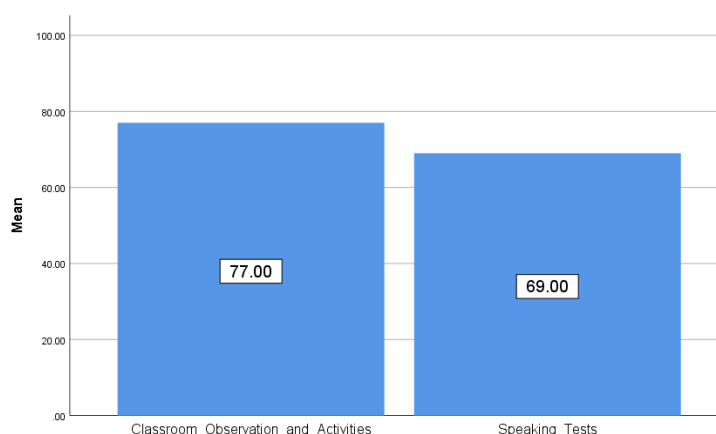


Figure 3. Percentages of assessment methods used by the teachers to test speaking

The analysis of the teachers' views on speaking tests indicates that most of the teachers focused on communicative skills in speaking assessments, with 66% frequently designing speaking tests that prioritized communication. The item with the highest mean score (4.19) indicates that clear instructions were provided for speaking tasks. The mean scores also reflect that the teachers often assessed language fluency and pronunciation ($M = 3.83$ and 3.85 , respectively). However, there were challenges regarding the inclusion of activities that promote fluency ($M = 3.68$), with 42% of the teachers indicating that they rarely or sometimes incorporated such activities.

The teachers also emphasized the importance of providing feedback on students' pronunciation and intonation, with 64% doing it so regularly. While the majority of the teachers felt confident in their ability to design effective speaking tests ($M = 3.75$), challenges were identified related to creating a supportive environment and providing opportunities for practice in different contexts ($M = 3.72$). Table 3 presents the teachers' views on their assessment practices for speaking.

Table 3. Teachers' views on designing speaking tests

Context	Mean	Never (%)	Rarely (%)	Sometimes (%)	Often (%)	Always (%)	Chi-Square	df	Sig.
1. I design speaking tests that focus on communicative skills.	3.86	0	4	30	42	24	30.24	3	0.000
2. I make sure my speaking tests cover the learning goals from the curriculum.	4.02	1	3	21	43	32	66.20	4	0.000
3. I provide clear instructions for each speaking task.	4.19	0	3	16	40	41	41.84	3	0.000
4. I include activities that promote fluency in spoken language.	3.68	2	5	37	35	21	53.20	4	0.000
5. I use already-prepared tests from different sources to assess students' speaking ability.	3.35	4	12	39	35	10	50.30	4	0.000
6. I ensure the suitability of prepared tests before using them.	3.90	2	6	24	36	32	46.80	4	0.000
7. I assess content in students' speaking.	3.64	1	9	37	31	22	44.80	4	0.000
8. I assess language fluency in students' speaking.	3.83	0	3	31	46	20	39.44	3	0.000
9. I provide engaging topics for speaking tasks in my tests.	3.72	1	3	39	37	20	65.00	4	0.000
10. I provide feedback on students' pronunciation and intonation.	3.85	3	2	31	35	29	52.00	4	0.000
11. I create a supportive environment where students feel comfortable speaking English.	3.91	1	6	25	37	31	49.60	4	0.000
12. I provide opportunities for students to practice speaking in different contexts.	3.72	0	8	31	42	19	26.00	3	0.000
13. I feel confident in my ability to design effective speaking tests.	3.75	1	7	30	40	22	51.70	4	0.000

A chi-square goodness of fit test was conducted to find out whether the respondents had a similar preference for the different options. The results, as displayed in Tables 2 and 3, showed significant differences among all the items ($p < 0.05$).

4.2. Analysis of the interviews with the teachers

The second phase of the study was qualitative in nature, where semi-structured interviews were conducted with 10 EFL teachers to answer the third research question of the study. The interviews provided deeper insights into the challenges teachers face in designing writing and speaking assessments. The interviews offered explanations for some of the trends observed in the quantitative data, particularly regarding the practical constraints they face in implementing curriculum goals.

4.2.1. Challenges in writing test design

Through the interviews, several challenges in designing writing tests were identified. The most frequently mentioned challenge was time constraints, as the teachers noted that grading papers and ensuring that test items were appropriate for students' levels were time-consuming. One teacher commented: *"We have to review every single essay and give feedback, which takes hours. But we also have classes to teach and other responsibilities. There just isn't enough time"*. The teachers also reported difficulties in balancing question difficulty with students' abilities and ensuring clarity in the questions. As another teacher explained, *"Ambiguity in test questions can lead to misinterpretation, which is why clarity is so important"*.

Other challenges included aligning questions with curriculum standards and ensuring diversity in question types. The teachers emphasized the need for varied formats, such as multiple-choice and fill-in-the-blank tasks, to assess different language skills. Despite these challenges, many teachers acknowledged the importance of providing feedback on students' writing to improve their language proficiency.

4.2.2. Challenges in speaking test design

Similar to writing assessments, time constraints were also a significant challenge for speaking tests. The teachers reported that large class sizes made it difficult to provide individual speaking practice and feedback. One teacher noted, "*There is not enough time to assess each student's speaking skills thoroughly, especially in large classes*".

Additionally, the teachers highlighted the difficulty in assessing spoken language reliably, particularly in creating test items that accurately reflect real-life communication. The teachers also discussed the challenges of balancing content coverage with the depth of students' understanding during speaking assessments. Lastly, resource constraints were mentioned by some participants. "*Resource limitations make it hard to incorporate interactive speaking activities like role-plays*". Despite these challenges, the teachers emphasized the importance of creating a supportive environment where students feel comfortable speaking and practicing their language skills. The summary is provided in Table 4.

Table 4. Challenges in designing assessments for writing and speaking skills

Skill	Challenges
Writing	Time constraints and effort, Balancing question difficulty and student levels, Ambiguity in questions, Matching questions to curriculum and standards, Diversity in question types
Speaking	Time constraints and limited practice, Assessment difficulty and developing reliable methods, Designing appropriate test items, Balancing content coverage and depth of understanding, Resource constraints

4.2.3. Teacher professional development

In terms of professional development, a significant number of teachers indicated that they felt the need for further training on designing effective writing and speaking assessments. Many teachers expressed a desire to attend workshops and read academic articles to improve their assessment practices. One teacher mentioned, "*Reading books and articles and attending workshops have proven to be highly beneficial in improving my assessment techniques*". Another added: "*We need more in-service training and workshops focused specifically on how to assess speaking and writing*".

V. DISCUSSION

The primary objective of this study was to explore how Iranian junior high school EFL teachers design writing and speaking assessments and to examine the challenges they face in aligning these assessments with the new curriculum. The integration of quantitative and qualitative findings provided a comprehensive understanding of teachers' practices, showing not only what patterns existed in their assessment design but also why these patterns occurred within the Iranian educational context. The revised national curriculum places strong emphasis on communicative competence as the central goal of English language instruction (Riahipour, Tavakoli, & Eslami Rasekh, 2019). Regarding assessment, it encourages the use of authentic, performance-based tasks that reflect real-life language use, especially in speaking and writing (Razavipour & Rezagah, 2018). Teachers are expected to move beyond traditional grammar-based tests and instead design assessments that evaluate learners' ability to convey meaning clearly, appropriately, and fluently in context. These principles reflect a shift toward CLT and require assessments to promote interaction, creativity, and meaningful communication. The study revealed a complex interplay between teachers' perceptions, their assessment practices, and the limitations imposed by the educational context.

The findings indicate that, while there has been a shift toward emphasizing communicative competence in both writing and speaking assessments, traditional assessment methods remain dominant in practice. In writing assessments, the preference for short-answer tasks (79%) and paragraph writing (58%) underscores the importance placed on evaluating grammatical accuracy and written expression. These findings agree with prior research suggesting that objective assessments like short-answer tasks are favored for their ability to yield clear and measurable results (Fitria, 2024; Shaban, 2021). However, the

moderate use of more communicative tasks such as picture-cued writing (45%) suggests that teachers struggle to fully integrate the new curriculum's emphasis on real-life communication skills into their assessments. This finding was further explained in the interviews, where the teachers reported that designing communicative tasks was often constrained by time, large class sizes, and heavy workloads. Several teachers mentioned that assessing open-ended writing tasks was especially time-consuming, which discouraged them from using such formats frequently.

One of the key challenges identified in the study was the balancing of question difficulty with students' proficiency levels. The teachers expressed frustration with the time constraints they face, which impact their ability to create fair and well-aligned assessments. This challenge is consistent with the findings from [Nguyen and Le \(2021\)](#), which highlighted the time-intensive nature of test design and grading as significant obstacles for teachers. The limited focus on communicative skills in writing assessments may also be attributed to the practical constraints teachers face in terms of time, resources, and training. This limitation emphasizes the need for more targeted professional development opportunities that help teachers design assessments that reflect the communicative goals of the curriculum.

Similarly, in speaking assessments, the study found a strong preference for interactive and realistic assessment methods, such as the question-and-answer method (77%) and role-plays (74%). These findings support the growing trend towards authentic assessment practices that focus on real-world language use, as discussed by [Gulikers et al. \(2004\)](#). The use of interactive assessments reflects the desire to assess students' ability to use language in context, which is a central component of the communicative approach promoted by the new curriculum. However, despite the preference for these interactive methods, the teachers faced significant challenges in assessing speaking skills. The interviews provided deeper insight into these difficulties. The teachers described assessing speaking as "the hardest part" of their job. They reported facing multiple challenges, including limited time, large class sizes, the difficulty of creating reliable and authentic assessment tasks, and the absence of standardized rubrics. These contextual factors help explain the moderate mean scores in the quantitative data for items related to fluency and feedback. Designing appropriate speaking test items, such as role-plays and interviews, is not easy, as it is hard to ensure that the tasks truly represent real-life communication. This view reflects what [De Jong \(2023\)](#) believes; he argues that assessing spoken language is inherently complex due to its transient, dynamic, and context-dependent nature.

The study also revealed that a lack of training in new assessment methods and a reliance on traditional assessment techniques are major barriers for teachers. Despite the emphasis on communicative competence in the curriculum, many teachers continue to use methods that focus primarily on grammar and vocabulary. This is in line with the findings of [Farhady and Tavassoli \(2021\)](#), who observed that Iranian EFL teachers were slow to adopt newer, more communicative assessment techniques due to insufficient training and a lack of familiarity with modern assessment trends. This gap between the curriculum's goals and teachers' practices suggests that more comprehensive and continuous professional development programs are needed to equip teachers with the skills and knowledge to design assessments that reflect the principles of communicative language teaching.

Furthermore, the teachers expressed concerns about the lack of appropriate resources and the large class sizes, which hinder their ability to implement more interactive and communicative assessments. These practical constraints are consistent with the findings by [Aliakbari et al. \(2023\)](#), who report that Iranian EFL teachers often face substantial difficulties when attempting to apply communicative assessment methods in overcrowded and under-resourced classrooms. In particular, the challenges of managing speaking assessments in large classes were emphasized, as they require time, individual attention, and suitable technological tools for tasks such as recording, evaluating, and providing feedback. The teachers in this study echoed the need for institutional support and access to relevant resources in order to implement assessment practices that conform more closely to the communicative goals outlined in the national curriculum. A combination of logistical limitations and administrative workload continues to be a major barrier to effective assessment reform in the Iranian junior high school context.

VI. CONCLUSION AND IMPLICATIONS

This study explored how Iranian junior high school EFL teachers design writing and speaking language tests under a new curriculum with a focus on communicative competence. The results highlighted the significant efforts by teachers to align their assessments with the curriculum's communicative goals.

However, the study also revealed several challenges that hinder the effective attainment of these objectives.

In speaking assessment, the teachers tended to rely on formats such as role plays and question-and-answer activities. While these approaches partially support communicative goals, they often fall short of evaluating interactional strategies, extended discourse, and spontaneous speech. Similarly, writing assessments are typically structured around short-answer or controlled paragraph writing tasks, which do not adequately measure learners' ability to convey meaning in authentic written contexts. These limitations suggest that many assessments prioritize practicality and clarity over communicative depth. These patterns point to a deeper issue, a gap between the curriculum's goals and teachers' assessment literacy. Teachers lack adequate support to design performance-based tasks that capture both fluency and meaningful language use. Addressing this disconnection requires targeted training focused on practical skills for assessment design. Training programs should include methods of developing open-ended writing prompts, structuring speaking tasks that elicit extended language, and using analytic rubrics to evaluate content, coherence, fluency, and interaction. Time management strategies and classroom-based solutions such as peer assessment and speaking rotations can also support the teachers working in large classes (Li, Huang, & Cheng, 2022).

Such professional development can be designed and delivered by local teacher training centers in collaboration with applied linguists and curriculum experts. Programs can combine theory with hands-on application and include classroom-based workshops, collaborative task design, and ongoing feedback cycles. These initiatives can be supported through regular in-service training days, school-based mentoring groups, and online platforms offering digital resources and examples of communicative assessments.

Another important avenue is the integration of educational technology into assessment practice. Teachers need access to tools that support efficient and communicative evaluation, such as mobile apps for recording speaking tasks, software for automated feedback on writing, and platforms that allow students to engage in interactive language production. These tools can make assessments more dynamic, manageable, and consistent with communicative goals, especially in settings where time and resources are limited.

In conclusion, the study reveals that, although Iranian junior high school EFL teachers make commendable efforts to adapt their test design to the communicative curriculum, they are constrained by systemic issues such as limited training, inadequate resources, and large class sizes. To bridge the gap between curriculum expectations and classroom realities, there is a pressing need for structured, context-sensitive professional development and better access to both human and technological support.

The findings have practical implications for policymakers, curriculum designers, and school administrators. Policymakers can invest in ongoing, context-sensitive professional development to strengthen teachers' skills in speaking and writing assessment. Curriculum designers can provide clear, communicatively aligned guidelines and adaptable examples of performance-based tasks and rubrics. School administrators can create supportive environments by allowing time for collaborative planning, facilitating access to digital tools, and reducing excessive workloads. Coordinated efforts across these stakeholders also help bridge the gap between curriculum goals and classroom assessment, promoting more authentic evaluation of students' communicative competence.

There are a number of limitations to this study that need to be acknowledged. Firstly, the study relied on convenience and purposive sampling methods, which, while practical, may introduce biases. As a result, the findings may not be fully generalizable to wider populations of Iranian teachers. Expanding the sample size and including teachers from diverse regions, school types, and varying levels of experience would provide a more comprehensive understanding of test design practices across different contexts in Iran. Secondly, the reliance on self-reported data through questionnaires and interviews may



introduce subjectivity and social desirability bias. Future studies could employ a longitudinal design to track changes in teachers' practices over time. Last but not the least, research could focus on analyzing actual test papers to verify self-reported data, providing more objective evidence of alignment with curriculum goals.

AUTHORS' BIOGRAPHIES

Firooze Foroozan is a PhD candidate in Applied Linguistics at Yazd University, Iran. Her areas of interest include language assessment, second language acquisition, and educational technology.

Mohammad Javad Rezai holds a PhD in Applied Linguistics from the University of Essex, England and is a faculty member of English Department, Yazd University, Iran. His research areas include language assessment, second language acquisition, and psycholinguistics.

REFERENCES

- Aliakbari, M., Yasini, A., & Sadeghi, S. (2023). Iranian EFL teachers' classroom assessment practices: Discrepancy between theory and practice. *International Journal of Language Testing*, 13(2), 149-169.
- Atai, M. R., & Mazlum, F. (2013). English language teaching curriculum in Iran: Planning and practice. *The Curriculum Journal*, 24(3), 389-411.
- Bartlett, M. S. (1954). A note on the multiplying factors for various χ^2 approximations. *Journal of the Royal Statistical Society. Series B (Methodological)*, 296-298.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101.
- Brown, H. D., & Abeywickrama, P. (2010). *Language assessment: Principles and classroom practices* (2nd ed.). Pearson Education.
- Bryman, A. (2006). Integrating quantitative and qualitative research: How is it done?. *Qualitative Research*, 6(1), 97-113.
- Cattell, R. B. (1966). The meaning and strategic use of factor analysis. In *Handbook of multivariate experimental psychology* (pp. 131-203). Boston, MA: Springer US.
- Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage Publications.
- Creswell, J. W. (2013). *Steps in conducting a scholarly mixed methods study*. DERB Speaker Series.
- De Jong, N. H. (2023). Assessing second language speaking proficiency. *Annual Review of Linguistics*, 9(1), 541-560.
- Farhady, H., & Tavassoli, K. (2021). EFL teachers' perceptions and practices of their language assessment knowledge. *Language Testing in Asia*, 11(1), 1-19.
- Fitria, T. N. (2024). Choosing objective and non-objective test as instrument for assessment of EFL student learning outcomes. *Globish: An English-Indonesian Journal for English, Education, and Culture*, 13(1), 1-19.
- Fulcher, G. (2000). The 'communicative' legacy in language testing. *System*, 28(4), 483-497.

- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language assessment quarterly*, 9(2), 113-132.
- Fulcher, G. (2015). *Re-examining language testing: A philosophical and historical perspective*. Routledge.
- Gulikers, J. T., Bastiaens, T. J., & Kirschner, P. A. (2004). A five-dimensional framework for authentic assessment. *Educational technology research and development*, 52(3), 67-86.
- Hinton, P. R., McMurray, I., & Brownlow, C. (2014). *SPSS explained*. Routledge.
- Inbar-Lourie, O. (2016). Language assessment literacy. In *Language testing and assessment* (pp. 1-14). Springer, Cham.
- Kaçar, K. (2023). Commonly used techniques in testing foreign language skills and communicative language teaching. *Sinerji Uluslararası Alan Eğitimi Araştırmaları Dergisi*, 4(2), 68-85.
- Kaiser, K. L. (1974). On the optical activity of polychlorinated biphenyls. *Environmental Pollution* (1970), 7(2), 93-101.
- Kardoust, A., & Saeedian, A. (2021). Iranian EFL teachers' conceptions and practices of communicative language teaching curriculum. *Issues in Language Teaching*, 10(2), 171-202.
- Karizak, A. F., Afraz, S., & Samimi, F. (2024). Designing and validating a self-assessment tool for improving the speaking skill of Iranian EFL learners. *Journal of New Trends in English Language Learning (JNTLL)*, 3(3).
- Khadrir-Sharabyan, S., Kheir-Abadi, R., Alavi-Moghaddam, S. B., Anani-Sarab, M. R., Forouzandeh-Shahraki, E., Ghorbani, N. (2014). *Prospect 1: Teachers' Guide*. Sherkate Chap va Nashre Ketabhaye Darsiye.
- Lantolf, J. P., & Poehner, M. E. (2014). *Sociocultural theory and the pedagogical imperative in L2 education: Vygotskian praxis and the research/practice divide*. Routledge.
- Li, J., Huang, J., & Cheng, S. (2022). The reliability, effectiveness, and benefits of peer assessment in college EFL speaking classrooms: Student and teacher perspectives. *Studies in Educational Evaluation*, 72, 101-120.
- Mellati, M., & Khademi, M. (2018). Exploring teachers' assessment literacy: Impact on learners' writing achievements and implications for teacher development. *Australian Journal of Teacher Education*, 43(6), 1-18.
- Namaziandost, E., Hashemifardnia, A., & Shafiee, S. (2019). The impact of opinion-gap, reasoning-gap, and information-gap tasks on EFL learners' speaking fluency. *Cogent Social Sciences*, 5(1), 1630150.
- Nevisi, R. B., & Hosseinpour, R. M. (2022). Task-based speaking assessment in an EFL academic context: A case of summative and formative assessment. *Research in English Language Pedagogy (RELP)*, 10(2), 256-276.
- Nguyen, T. M., & Le, G. N. H. (2021). The influence of COVID-19 stress on psychological well-being among Vietnamese adults: The role of self-compassion and gratitude. *Traumatology*, 27(1), 86-97.
- Noughabi, M. A. (2017). Current pedagogical challenges in Iranian EFL teachers' views: A qualitative study. *Journal of Education and Practice*, 8(9), 217-228.



- Popham, W. J. (2011). *Assessment literacy overlooked: A teacher educator's confession*. *The Teacher Educator*, 46(4), 265-273.
- Razavipour, K., & Rezagah, K. (2018). Language assessment in the new English curriculum in Iran: Managerial, institutional, and professional barriers. *Language Testing in Asia*, 8(1), 1-18.
- Riahipour, P., Tavakoli, M., & Eslami Rasekh, A. (2019). A Change without a change!: National curriculum reform and EFL teachers' motivation in Iran. *Journal of English Language Teaching and Learning*, 11(23), 293-325.
- Riazi, A. M., & Razavipour, K. (2011). (In) Agency of EFL teachers under the negative backwash effect of centralized tests. *International Journal of Language Studies*, 5(2) 122-142.
- Sa'diyah, A. (2020). Alternative assessment practices and difficulties on EFL students' speaking skill. In *International Joint Conference on Arts and Humanities (IJCAH 2020)* (pp. 810-815). Atlantis Press.
- Sadeghi, K., & Richards, J. C. (2016). The idea of English in Iran: An example from Urmia. *Journal of Multilingual and Multicultural Development*, 37(4), 419-434.
- Scarino, A. (2013). Language assessment literacy as self-awareness: Understanding the role of interpretation in assessment and in teacher learning. *Language Testing*, 30(3), 309-327.
- Shaban, A. M. S. (2014). A comparison between objective and subjective tests. *Journal of the College of Languages (JCL)*, 30, 44-52.
- Stiggins, R. J. (1991). Assessment literacy. *Phi Delta Kappan*, 72(7), 534-539.
- Stratton, S. J. (2021). Population research: convenience sampling strategies. *Prehospital and disaster Medicine*, 36(4), 373-374.
- Taylor, L. (2013). Communicating the theory, practice and principles of language testing to test stakeholders: Some reflections. *Language testing*, 30(3), 403-412.
- Tayyebi, M., Abbasabady, M. M., & Abbassian, G. R. (2022). Examining classroom writing assessment literacy: A focus on in-service EFL teachers in Iran. *Language Testing in Asia*, 12(1), 1-20.
- Vogt, K., & Tsagari, D. (2014). Assessment literacy of foreign language teachers: Findings of a European study. *Language Assessment Quarterly*, 11(4), 374-402.
- Xu, Y., & Brown, G. T. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, 58, 149-162.
- Yuzar, E. (2020). Incorporating communicative competence in assessment and English language teaching in multilingual settings. *REiLA: Journal of Research and Innovation in Language*, 2(1), 8-13.
- Zayniddinova, R. (2025). The grammar-translation method in language teaching: Educator insights. *Journal of Teaching Theory and Practice*, 1(3), 169-176.