

EXPERIENCED VS NOVICE EFL TEACHERS: A STUDY OF SPEAKING ASSESSMENT LITERACY

Cansu Aykut Kolay¹, Büşra Ulu²

Department of Foreign Language Education, Kocaeli University, Turkey
English Preparatory School, Gebze Technical University, Turkey
E-mail: cansu.aykut@kocaeli.edu.tr

Received: 2024-11-06

Accepted: 2024-12-08

Published: 2024-12-20

Abstract

Assessment is a core component of foreign language education, significantly influencing teaching and learning processes. Since various skills, including reading, writing, listening, speaking, and grammar, must be assessed separately, language teachers need to be assessment literate in each skill. Among these, Language Assessment Literacy (LAL) is critical for effective teaching. While numerous studies have explored teachers' assessment literacy for reading, writing, listening, and grammar, limited research focuses on speaking assessment literacy. This study addresses this gap by examining the speaking assessment literacy of novice and experienced teachers. Using a qualitative research design, data were collected from 16 EFL teachers (eight novice and eight experienced) through nine semi-structured interview questions adapted from established question repositories. Thematic analysis was employed to examine the data across three aspects: knowledge, beliefs, and practices. The findings highlighted differences and similarities between the two groups. While experienced and novice teachers shared common speaking assessment practices, differences emerged in their overall levels of speaking assessment literacy. The results underscore the need for tailored pre-service and in-service training programs focusing on skill-based language assessment. These programs should be integrated into teacher education and induction to enhance assessment competence across varying levels of teaching experience. This study contributes to the understanding of speaking assessment literacy, offering insights for improving teacher training in foreign language education.

Keywords: *assessment, assessment literacy, speaking, speaking assessment*

1. Introduction

Assessment literacy (AL) has been the focus of interest in the field of English Language Teaching (ELT) especially over the last decade. The time when AL gained importance started with the introduction of "The Standards for Teacher Competence in Educational Assessment of Students" (1990). Coined by Stiggins in 1991, AL was defined by several researchers by focusing on different aspects such as skills and knowledge. AL is about having competence in assessment of a certain area like language, having necessary knowledge and applying it into practice of assessment (Fulcher, 2012; Stiggins, 1995). As

insufficient assessment has a negative impact on students' learning and institutions' well-being, it is essential for teachers to be literate in terms of the issue of assessment.

As a sub-branch of the concept of AL, language assessment literacy (LAL) has been focused and defined by numerous researchers (Davies, 2008; Inbar-Lourie, 2017; O'Loughlin, 2013; Pill & Harding, 2013; Vogt & Tsagari, 2014). It has been argued that literacy in language assessment has a number of aspects that require EFL teachers to be competent in order to meet not only their students or institutions' expectations but also their own ones as well. In this context, there are a bunch of studies conducted in the international context to measure EFL teachers' language assessment literacy. (Crusan et al., 2016; Jannati, 2015; Kiomrs et al., 2011; Mertler, 2003; Montee et al., 2013; Muñoz et al., 2012; Tajeddin et al., 2018; Shim, 2009). Overall, it was concluded that EFL teachers need training in assessment literacy and there is a mismatch between their theoretical beliefs and practices about language assessment.

As for the Turkish context, the related literature includes quite a few studies conducted in the field of LAL (Büyükkaracı, 2016; Gök, Erdogan, & Altinkaynak, 2012; Mede & Atay, 2017; Ölmezer-Öztürk, & Aydın, 2018; Sarıyıldız, 2018; Tamerer, 2019), which creates a gap in the related literature. Building on previous studies on language assessment literacy in the Turkish context, this study aims to investigate the knowledge, beliefs, and practices of experienced and novice EFL teachers regarding speaking assessment at the university level in Turkey. Additionally, it seeks to explore how teaching experience influences their perceptions and practices of speaking assessment. To achieve these objectives, the study addresses the following research questions:

1. What are the beliefs and practices of EFL teachers concerning speaking assessment literacy?
2. How do novice and experienced EFL teachers differ in their speaking assessment literacy?

2. Literature Review

2.1 Assessment Literacy (AL)

Assessment has been of significance in the field of teaching as it is the indicator of the extent to which learners grasp from the input they are provided. Assessment literacy has gained importance after the publication of the document *The Standards for Teacher Competence in Educational Assessment of Students* (1990). Stiggins (1995) explains assessment literacy is about knowing what it is assessed, why it is assessed, how to representatively assess it, what problems can emerge during the process of assessment, and how to prevent those. In other words, teachers who are assessment literate know what to assess, how and why they assess, and they can foresee possible problems and have the vision to prevent them before they occur. A comprehensive definition of assessment literacy in terms of its functionality is suggested by Fulcher (2012):

“The knowledge, skills and abilities required to design, develop, maintain or evaluate, large-scale standardized and/or classroom-based tests, familiarity with test processes, and awareness of principles and concepts that guide and underpin practice, including ethics and codes of practice. The ability to place knowledge, skills, processes, principles and concepts within wider historical, social, political and philosophical frameworks in order understand why practices have arisen as they have, and to evaluate the role and impact of testing on society, institutions, and individuals.” (p. 125).

In terms of the significance of assessment literacy, Koh et al. (2017) claimed that it is essential for teachers to know how to make use of large-scale tests and assess them, but it is also necessary to have qualifications to design and use good quality classroom assessment tools which enhance and support learning. Therefore, higher levels of teacher assessment literacy are also conducive to students' learning. Giraldo (2018) states assessment literacy enlarged teachers' competence at monitoring, recording, improving, and reporting on the process of students' learning. With respect to this, he indicates that according to Standards for Teacher Competence in Educational Assessment of Students (AFT, NCME, & NEA, 1990) there are two factors which teachers should take into consideration while assessing learners: The first one is teachers' ability to choose, design, and evaluate valid assessments for positive effects on learning, teaching, and schools. The latter is the appropriate use of tests and test results. However, when it comes to the application of those, it is suggested that today's teachers' assessment literacy level is not satisfactory (Popham, 2010). Correspondingly, studies conducted by Wiggins and Pincus (1989) and Bol et al. (1998) revealed many teachers are insufficiently trained and poorly prepared to develop, administer and interpret the outcomes of various assessment types. Besides, Stiggins (2010) points out that teachers' level of assessment literacy may be lower because of their limited exposure programs of professional development, which may be worth investigating with regard to language teachers' lower levels of assessment literacy.

2.2 Language Assessment Literacy (LAL)

As language teachers' level of assessment literacy has a huge impact on the quality of education, it has been highly investigated and has become a focus of great interest in existing body of the literature on language assessment literacy (LAL). There is a considerable number of studies examining the language teachers' knowledge and practices of assessment. Shim (2009) conducted a study investigating teachers' beliefs and practices regarding English language assessment by using questionnaire and interviews with some of the participants teachers. The findings of the study revealed that whereas the English teachers in the study were assessment literate and knowledgeable about the tenets of assessment and testing, they did not make use of these principles in their actual classroom practices. In their study with 62 Columbian teachers in a private university, Muñoz et al. (2012) demonstrated that there is a mismatch between teachers' beliefs and classroom practices of assessment. They suggested that teachers need opportunities to reflect on their practice of assessment and more guidance on formative assessment tools. Kiomrs et al. (2011) found out that Iranian EFL teachers' level of assessment literacy is not satisfactory. Correspondingly, Montee et al. (2013) concluded that only a small number of participants who had taken a college course in language assessment felt more confident in their assessment practices when compared to the majority who had not taken any assessment course beforehand, which decreased their level of assessment literacy.

Pill and Harding (2013) put forward that just like many other 'new' literacies such as computer, literacy and scientific literacy, LAL is particularly associated with reading and writing skills. Within the framework of AL, LAL is also fundamental. Inbar-Lourie (2017) claimed that LAL has drawn considerably from the literature and research on general AL, while attempting to set itself apart as a knowledge base that incorporates unique aspects inherent in theorizing and assessing language-related performance. In other words, language as a construct for assessment is the major difference between LAL and general AL. In this regard, LAL has been defined by several researchers. AL is described as the ability to design,

develop and evaluate tools of assessment, together with the ability to observe, evaluate, and score assessments based on theoretical knowledgebase (Vogt & Tzagari, 2014). Davies (2008) also asserted that language assessment literacy has three dimensions, *skills*, *knowledge*, and *principles* respectively. The aspect of skills includes suitable methods like item writing, using statistics, and analysis of tests. Correspondingly, knowledge requires familiarity with measurement tools and awareness of language items being tested. Lastly, principles are based on the appropriate use of language tests and being cautious about their fairness and the issue of wash back (Davies, 2008). Additionally, Pill and Harding (2013) pointed out that LAL encompasses several competencies that allow the teachers to comprehend, evaluate and come up with language tests and analyze test data. Finally, O'Loughlin (2013) suggests LAL includes a range of skills related to production of tests, score interpretation, and test evaluation in line with the development of an insight into functions and roles of assessment in society. As these given definitions suggest, language teachers need to be aware of the processes of constructing and analyzing test items, conducting the tests, grading, and evaluation of grades by using statistical tools based on their background theoretical knowledge and principles of assessment.

2.3 Speaking Assessment Literacy

Tajeddin et al., (2018) examined the speaking assessment literacy of experienced and novice EFL teachers. To assess participant EFL teachers' speaking assessment literacy, they came up with three different scenarios in which they are supposed to measure their learners' speaking skill. Besides, they asked interview questions like "*How do you describe the purpose of speaking assessment in your class?*", "*What do you assess mostly in your learners' speaking performances?*" and "*What tasks do you use to assess your learners' speaking?*". That study is of significance because it is among a few studies in the existing body of literature focusing on skill-based assessment literacy between novice and experienced teachers. When it comes to the results of that study, they found out that summative content analysis resulted in three main categories of assessment purposes: *formative assessment*, *diagnostic assessment*, and *summative assessment*. Likewise, they concluded that teachers took into consideration accuracy, fluency, vocabulary, pronunciation, communication, interaction, pragmatics, and content while assessing speaking skill. In terms of the assessment methods, it was shown that they made use of methods like role plays, monologues, and authentic tasks. Overall, they came up with the conclusion that novice and experienced teachers have both common and divergent speaking assessment literacy patterns. In an earlier study conducted by Rubin and Mead (1985), it was stated that the assessment of speaking skill was done based on two approaches called observational and structured respectively. In the former one, students' natural behaviors are observed without intervention. On the other hand, students are supposed to work on a specific oral task then, their performance is evaluated. They claimed that an authentic context for communication is a key factor in both approaches, which is also confirmed by Zaim et al. (2020) stating that role-play, picture description and dialogues are the most preferred speaking assessment tools. As well as the most common ones, other employed tools are storytelling, presentation and interviews (Zaim et al., 2020).

2.4 Experienced and Novice Teachers' Level of Language Assessment Literacy

In addition, there have been studies investigating the effect of experience on assessment literacy of teachers. Based on the results of the study, Popham (2011) stated

that experienced teachers are less likely to help teachers who have started the job as they lack knowledge of testing despite their experience. Parallely, Mertler (2003) also investigated the role of classroom experience on assessment literacy of pre-service and in-service teachers. Data were collected through Classroom Assessment Literacy Inventory (CALI) which was developed for this purpose. The findings of the study revealed that the experience of teachers has an impact on assessment literacy and there were significant differences between in-service teachers and pre-service teachers. However, Jannati (2015) studied the extent to which EFL instructors reflect their assessment literacy to their practices and whether experience of teacher has an impact on assessment literacy and practices of teachers. The findings showed that English instructors were assessment literate, but their perceptions of assessment literacy were not reflected in their practices. Regarding experience it was found that experience does not influence assessment literacy.

Considering the Turkish context, the number of studies conducted in the field of language assessment literacy tends to incline. Büyükkarcı (2016) concluded that in-service English teachers' level of assessment literacy is not satisfactory no matter what their teaching level is. Therefore, further training in language assessment is in favor of these language teachers.

In the study done by Gök et al. (2012), it was stated that the majority of the teachers participated in their study had negative perceptions about assessment. For instance, they labelled the process of assessment as 'numerical data', which created a negative connotation towards assessment. Besides, Mede and Atay (2017) found out that EFL teachers have insufficient training in the areas like preparing classroom tests and providing feedback. Another study by Atay and Öz (2015) was conducted with both experienced and novice preparatory school English teachers through semi-structured interviews. The results indicated that teachers were aware of basic classroom-based assessment principles and tools, but there was a discrepancy between their beliefs and practices. It was revealed that experience was not an important factor for beliefs about assessment.

Regarding skill-based assessment literacy, Ölmezer-Öztürk and Aydın (2018) conducted a study with 542 teachers through a survey based on asking beliefs of four constructs: writing, speaking, reading and listening. The findings revealed that assessment literacy of teachers was not at an optimal level. It was also found that the construct in which the teachers were the most knowledgeable was reading assessment while the lowest one was the assessment of listening.

Another study with regard to skill-based assessment literacy has been conducted by Genç et al. (2019). In the study, 58 high school EFL teachers' skill-based assessment literacy level was investigated within the Turkish context. Based on a quantitative design, it was found out that the participant's assessment of writing was lower but the one for speaking assessment was over when compared to their average score on the implemented scale.

Regarding the aforementioned background, it is apparent that although the field of general AL has been highly investigated, the number of studies regarding skill-based LAL is quite few in both international and local contexts. Particularly, studies based on speaking assessment literacy are thin on the ground. That is the reason why speaking assessment literacy was chosen as the focus of this study. Besides, while most studies are based on a quantitative design in which participants' level of skill-based language assessment literacy is measured through a scale, the current study is based on a qualitative design so that thicker and deeper data can be reached and analyzed thematically. In this vein, this study addresses the gap in the existing body of literature regarding discrepancies between experienced and

novice EFL teachers in terms of their levels of speaking assessment literacy and their different preferences to measure students' speaking proficiencies in EFL classes.

3. Research Method

3.1 Participants

The participants of this study were 16 EFL teachers who are working as English instructors at foundational universities in different cities of Turkey including İstanbul and Antalya. For the purpose of the study, the teachers were divided into two groups based on their experience level. In the field of second language teaching, the concept of experience has been operationalized in various ways. Some researchers claim that any year of teaching is enough to consider teachers as experienced (Akyel, 1997; Mackey et al, 2004). McNeil (2005) states that two years of teaching is necessary to be labelled as experienced. In addition, a few several have the idea that teachers should have at least seven years of teaching to be called as experienced (Lopez, 1995); Gurzynski-Weiss, 2010). However, within the framework of this study, the researchers took three years as reference to define experienced teachers, which was in line with the operationalization of experience done by Mok (1994). Therefore, novice teachers' experience in the study was between zero and two years and that of experienced teachers was three and more years. Novice teachers' age ranges from 23 to 25 and experienced teachers' age range is between 26 and 41.

3.2 Instruments

A mixture of semi-structured interview questions adapted from Tajeddin, Alemi, and Yasaei (2018) and Öz and Atay (2015) were posed to the participants. There were nine questions in the interview, and they are particularly designed to find out EFL teachers' beliefs and practices of speaking assessment. The questions have been formulated in the light the opinions of three experts in the field of English Language Teaching and fine-tuned accordingly. The formulated questions are as follows:

- 1) How do you describe the purpose of speaking assessment?.
- 2) What is generally assessed in learners' speaking performances and what do you assess mostly in your learners' speaking performances?
- 3) What are the tasks used to assess speaking and which ones do you usually prefer?
- 4) What are the characteristics of a good speaking assessment?
- 5) Do you support the idea that instructors need to have some sort of background knowledge about assessing speaking?
- 6) Do you think students' speaking scores truly represent their speaking competency?
- 7) How do you make use of speaking assessment results?
- 8) To what extent are you convinced that your speaking assessment is valid/fair?
- 9) How do you increase your knowledge about speaking assessment? (e.g., reading books, attending workshops etc.)

3.3 Data collection

Data collection involved a combination of semi-structured interviews designed to explore participants' perspectives comprehensively. The interview questions were carefully developed to align with the research objectives and were adapted from established frameworks to ensure validity and relevance. Some of the interviews were conducted in face-to-face settings, providing an opportunity for direct interaction and richer data collection through the observation of non-verbal cues. Each face-to-face interview lasted

approximately 10–15 minutes, allowing sufficient time for participants to elaborate on their responses while maintaining their engagement. To accommodate varying schedules and preferences, additional interviews were conducted remotely using video conferencing platforms, ensuring flexibility and inclusivity. All interviews were audio-recorded with participants' consent to facilitate accurate transcription and analysis. The semi-structured format allowed for a balance between structured inquiry and open-ended responses, enabling participants to express their views freely while ensuring the consistency of core questions across interviews. This approach provided both depth and breadth in the data, capturing nuanced insights into participants' experiences, beliefs, and practices related to the research topic. The collected data was subsequently transcribed verbatim and subjected to rigorous qualitative analysis, ensuring that the findings were grounded in the authentic voices of the participants.

3.4 Data Analysis

A thematic analysis was conducted in the scope of the current study as it is regarded as an accessible, practical, and useful popular method among other qualitative data analysis techniques (Braun & Clarke, 2006). In this regard, the steps followed while conducting thematic analysis are:

1. Familiarizing with the collected data by re-reading it
2. Coding the relevant data
3. Searching for themes: reviewing the codes come up with considerably prevalent patterns
4. Revising emerging themes: checking them whether they need to be extracted or not
5. Making the definitions of the themes: selecting an appropriate name for each theme
6. Reporting: contextualizing the themes in line with the findings of the existing body of literature

As for the coding processes, the data were subjected to initial and second-level coding processes (Dörnyei, 2007) While the transcribed and cleaned data are read and emerging codes are detected within the context of initial-level coding, the coded data are re-read and unrecognized candidate themes are also reached through keyword search in second level coding.

With regard to the validity of the data analysis process, a process of interrater reliability was conducted which is of significance for the validation of thematic analysis procedure ensured (Mackey and Gass, 2005; Loewen & Philp, 2012). An external coder, who pursued her B.A. and M.A. degrees in the field of English Language Teaching, was assigned to re-code the data. 20% of the total interview transcriptions was given to the external rater which is regarded as an acceptable amount (Loewen & Philp, 2012). For each category (e.g., Knowledge, beliefs and practices), the interrater reliability was calculated separately. In this regard, the ratios reached of agreement were 98%, 100%, and 100% for each category, respectively. The percentages are acceptable since percentages above 75% are regarded as "good", but the ideal percentage is above 90% (Mackey & Gass, 2005, p.244).

3.5 Limitations

As for the limitations of the current study, the number of participants, and the lack of quantitative data can be mentioned. The data collected is only valid for the small number of participants which may violate the transferability of the results. Besides, triangulation of the

collected data is not possible as quantitative data were not collected within the context of the study.

4. Results and Discussion

4.1 Knowledge of speaking assessment

Table 1.

Background knowledge about speaking assessment

	Undergraduate	Frequency
Experienced	5	62.5
Novice	8	100
Total	13	81.25

81.25% of the total participants had speaking assessment courses in their undergraduate level. Because some of the participants came from different undergraduate backgrounds such as American literature and translation and interpreting studies, not all of them took courses related to assessment. In total, 62.5 of experienced teachers took an assessment course in their undergraduate level. Among five participants taking such courses in their undergraduate education, two of them stated that they benefited from them in terms of test material design while the other three took advantage of those courses regarding theories and testing different skills. When it comes to novice EFL teachers, all of them took an assessment course in their undergraduate program. Six of them indicated that they make use of the knowledge that they had during the process of preparing testing materials. In addition to this, two out of these six participants asserted that those courses made them more aware of concepts like reliability and validity of the test items. The rest two people stated that they used their undergraduate assessment knowledge while planning their lessons and testing students' language skills. As a comparison, it can be seen that novice teachers outnumber experienced teachers in terms of taking undergraduate assessment courses and the majority of them use that knowledge while preparing testing materials.

Table 2.

In service training

	In-service	Frequency
Experienced	5	62.5
Novice	3	37.5
Total	8	50

Considering in-service assessment training, the numbers of experienced teachers are higher than novice teachers with the percentage 62,5. Five novice teachers stated that they attended to in-service teacher training programs about assessment, but only three of them

said that they participated in speaking assessment workshops. On the other hand, 62,5% of experienced teachers took part in speaking assessment workshops at universities in which they are currently working. Both experienced and novice teachers specified that those speaking workshops were mostly related to standardization of speaking assessment.

When the participants were asked about how they improve their speaking assessment literacy skills, six categories emerged from the analysis as the ways to increase knowledge about speaking assessment which are workshops, reading articles and guidebooks, attending certificate programs and conferences, consulting expert knowledge and taking online courses. Majority of both novice and experienced teachers stated that they attended speaking assessment workshops to improve their knowledge. A minority of participants indicated that they also benefited from other categories as well. A comparison of the answers of both novice and experienced teachers revealed the same categories, showing no difference between two teacher groups in terms of preferences.

4.2 Beliefs on Speaking Assessment

Respondents were asked to describe the purpose of speaking assessment, and the common themes observed were namely, assessment of communicative competence and assessment of speaking skill. Dell Hymes (1972) stated that learners acquire competence as to when to speak, when not, and as to what to talk about with whom, when, where, in what manner. Whereas only one of the inexperienced teachers responded that the aim should be giving feedback to students' language improvement, the rest of them emphasized communicative competence as the main purpose of speaking assessment. However, the number of experienced teachers used the term "communicative competence" to describe the goal of assessment was lower than novice ones. Only three of them focused on the concept of communicative competence. Rather, they indicated that the main purpose of speaking assessment is to evaluate learners' mere speaking ability. They did not specify factors such as context, communication strategies or the role of interlocutors. Lastly, one of the experienced interviewees asserted that the goal is to get information about students "Language development".

In terms of the characteristics of good speaking assessment, a common view amongst novice interviewees was that evaluation should be based on rubrics which will assist in ensuring reliability, validity and objectivity. However, only one of the experienced teachers touched upon the use of rubric as a characteristic of good speaking assessment. Besides, novice teachers also stated that it is important to look at organization and coherence while assessing speaking, which was not observed in the responses of experienced teachers. One experienced teacher commented that it is important for teachers to have positive attitudes towards students while speaking to encourage and motivate students. Apart from these, both novice and experienced teachers focused on the importance of authentic task design to assess learners' accuracy, fluency, grammar and pronunciation skills.

All the experienced and novice interviewees of this study emphasized the necessity of the fact that teachers should have a background knowledge about speaking assessment with the responses such as "Definitely", "Yes, of course", "For sure" and "Definitely, yes". Some of the teachers added extra comments about why being competent at assessment of speaking is desirable. For instance, one experienced stated that "being an assessor requires different skills than being an instructor". In other words, teaching how to speak does not mean that teachers are competent enough to assess the speaking skill of students. One of

the novice teachers commented that teacher competency is important as language is about mostly communications and so speaking is of great importance. In addition, another comment from a novice teacher was related to validity and fairness of speaking assessment. Regarding the reliability and validity of learners' speaking assessment scores, seven of the inexperienced teachers commented that scores do not truly represent learners' speaking competency. There may be external and internal factors affecting learners' performance. Therefore, performance does not always show competence. Only one of them stated that it depends on the assessment type and evaluation criteria. Similarly, all experienced teachers in the study agreed that performance on an assessment task is highly unlikely to show how competent the students are and there are some other factors affecting students' performance including learners' affective filters like having anxiety. Correspondingly, another teacher commented that such true representation is not possible due to subjectivity even though rubrics are used.

Moreover, six of novice teachers suggested that rubrics need to be used to sustain fairness of the speaking assessments. Additionally, reliability and validity of test items should be ensured as well. One of them stated they he needs more training about scoring because he does not think that the scores that he has given are fair all the time. The general tendency among inexperienced teachers was that speaking assessment is quite valid and fair if there is a testing and assessment office in the institutions they work. As for experienced teachers, they mostly have the idea that their speaking assessments are fair as long as two raters' scores are parallel. Rather than rubrics, they particularly focused on the necessity of more than one rater in speaking assessments. Additionally, one of them suggested that the use of some scoring guides is helpful.

4.3 Practices on Speaking Assessment

Regarding the practice of speaking assessment, the participants were asked about what they generally assess in students' speaking when students are having a speaking exam or doing speaking tasks. Several different categories were observed. They were accuracy, fluency, vocabulary, communication described as comprehension and relevance of answers, pronunciation, content knowledge, coherence between ideas and finally delivery of speech including posture, eye contact and intonation.

Table 3.

What do teachers assess in a speaking task or speaking exam?

What is assessed?	Experienced %	Novice %
Accuracy	87.5	87.5
Fluency	87.5	75
Vocabulary (lexical knowledge)	75	62.5
Communication (appropriacy-comprehension- interaction)	37.5	75
Pronunciation	37.5	25
Content knowledge	25	25

Coherence	12.5	25
Delivery of speech (rate of speech-intonation-body language-eye contact)	12.5	25

In most of the categories identified significant differences were not found in all eight parts and at least one answer was given by both experienced and novice teacher. In terms of accuracy and content knowledge the numbers of teachers assessing those two aspects were equal. However, the number of experienced teachers teaching surpassed the number of novice ones in three categories: namely fluency, vocabulary and pronunciation. Contrastingly, the novice teachers outnumbered in communication, coherence and delivery of speech. In the categories in which the number of teachers differ, the differences were only one teacher and therefore it can be said that they were not significant except one. The only one was communication which included the answers such as “comprehension”, “appropriate answers” and “interaction” and in this category the number of novice teachers were six while for experienced it was three. This result was highly significant compared to the other seven criteria taken into consideration while assessing speaking.

Table 4.

The tasks used by novice teachers in speaking assessments

Task type	Number	Frequency (%)
Questions	6	75
Picture-description	1	12.5
Discussion- Debates	3	37.5
Project presentations	2	25
Role plays	2	25
Video or Voice-recording	1	12.5
Interviews	1	12.5

As can be seen, questions take up the most part in the tasks used by novice teachers in speaking assessments.

Table 6.

The tasks used by experienced teachers in speaking assessments

Task type	Number	Frequency (%)
Questions	5	62.5
Picture-description	2	25
Discussion- Debates	2	25

Project presentations	2	25
Role plays	0	0
Video or Voice-recording	0	0
Interviews	1	12.5

Generally, questions are the most common tasks used in speaking assessments by both novice and experienced teachers. Both groups of teachers stated that they prefer to ask open-ended opinion questions in which students can use their own background and imagination. After questions, debates and discussions are among the most preferred task types by teachers. They indicated that they generally opt for pair discussions or group debates. Picture descriptions, project presentations, and interviews were also found to be useful in speaking assessments. As one difference between novice and experienced teachers, while a small number of novice teachers benefit from role plays and video or voice-recording speaking assessment tasks, experienced teachers did not state such a practice. This may be because of the fact that the number of novice teachers having undergraduate training on speaking assessment is higher than the number of experienced teachers. In such undergraduate methodological courses, such kind of methods of speaking assessment is emphasized frequently as the participants suggested. This may be the reason why there is such a discrepancy between experienced and novice teachers in terms of selecting video-recordings or role plays in their speaking assessment, which is mainly driven by the theoretical background that they got in their undergraduate education. Another question posed to teachers was about how teachers make use of students' speaking scores. Six of experienced teachers responded that the score composed a certain percentage overall score of a student taking the speaking test and it was used to determine whether students have sufficient competency for the following level and whether they would pass. However, only two novice teachers talked about this issue. The other two teachers mentioned diagnostic goals. They commented that the scores of speaking assessments give information about students' needs, and they modify their speaking tasks based upon these needs identified with the scores. Contrastingly, the number of novice teachers who mentioned diagnostic goals outnumbered experienced teachers. Four novice teachers stated that speaking assessment results were used for diagnostic purposes. That is, students' weaknesses and strengths were detected, and necessary actions were taken in accordance with the results to correct the situation such as giving feedback. Lastly, two of inexperienced teachers was not able to give any relevant response as they do not assess speaking.

5. Conclusion

This study analyzed the data under three main categories: knowledge, belief, and practice. The findings revealed differences between experienced and novice teachers regarding their knowledge of speaking assessment. While novice teachers had exposure to assessment courses during their undergraduate education, this was less common among experienced teachers, many of whom gained knowledge through in-service training. However, not all experienced teachers had participated in such training, highlighting a gap in their formal preparation. Both groups emphasized the importance of background knowledge in ensuring fair, valid, and reliable assessments, aligning with prior research (Mede & Atay,

2017). Regarding beliefs, novice teachers prioritized communicative competence, considering factors like context, appropriacy, and communication strategies. This contrasts with experienced teachers, who placed less emphasis on these strategies, potentially due to differences in exposure to 21st-century teaching approaches (Piasecka, 2018). Novice teachers also favored using rubrics for their perceived role in ensuring validity and fairness, influenced by exposure to international frameworks like CEFR (Council of Europe, 2020). In contrast, experienced teachers preferred involving multiple raters, reflecting different assessment traditions. Both groups agreed that speaking scores might not always represent learners' true abilities, emphasizing the need for extended assessments to reduce anxiety. In practice, both groups frequently used open-ended questions, aligning with their belief in assessing communicative competence. However, novice teachers prioritized interaction and communication skills, while experienced teachers focused more on vocabulary, fluency, and accuracy. Teachers primarily used assessment results for diagnostic purposes and grading, but further leveraging these results to adapt lesson plans could enhance learning outcomes. This study contributes to understanding speaking assessment literacy by offering a qualitative, thematic analysis of teachers' perspectives. The findings underline the importance of skill-based assessment training in both undergraduate and in-service programs. Expanding such training can help bridge gaps between teachers' beliefs and practices, ensuring holistic and effective speaking assessments. Curriculum updates for undergraduate teacher education programs and workshops on assessment methods, particularly for experienced teachers, can enhance assessment literacy and foster positive washback in teaching and learning. Additionally, encouraging teachers to engage with academic literature on assessment literacy can deepen their knowledge and improve classroom practices.

References

- Akyel, A. (1997). Experienced and student EFL teachers' instructional thoughts and actions. *Canadian Modern Language Review*, 53, 678–704.
- American Federation of Teachers, National Council on Measurement in Education, & National Education Association. (1990). *Standards for teacher competence in educational assessment of students*.
- Bol, L., Stephenson, P. L., O'Connell, A. A., & Nunnery, J. A. (1998). Influence of experience, grade level, and subject area on teachers' assessment practices. *The Journal of Educational Research*, 91(6), 323–330.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
- Büyükkarçı, K. (n.d.). Identifying the areas for English language teacher development: A study of assessment literacy. *Pegem Eğitim ve Öğretim Dergisi*, 6(3), 333–346.
- Crusan, D., Plakans, L., & Gebril, A. (2016). Writing assessment literacy: Surveying second language teachers' knowledge, beliefs, and practices. *Assessing Writing*, 28, 43–56.
- Council of Europe. (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume*. Strasbourg: Council of Europe Publishing.
- Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing*, 25(3), 327–347.
- Dörnyei, Z. (2007). *Research methods in applied linguistics*. New York, NY: Oxford University Press.

- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113–132.
- Genç, E., Çalışkan, H., & Yüksel, D. (2020). Language assessment literacy level of EFL teachers: A focus on writing and speaking assessment knowledge of the teachers. *Sakarya University Journal of Education*, 10(2), 274–291.
- Giraldo, F. (2018). Language assessment literacy: Implications for language teachers. *Profile: Issues in Teachers' Professional Development*, 20(1), 179–195.
- Gök, B., Erdogan, O., Altinkaynak, S., & Erdogan, T. (2012). Investigation of pre-service teachers' perceptions about the concept of measurement and assessment through metaphor analysis. *Procedia-Social and Behavioral Sciences*, 46, 1997–2003.
- Gurzynski-Weiss, L. (2010). Factors influencing oral corrective feedback provision in the Spanish foreign language classroom: Investigating instructor native/nonnative speaker status, SLA education, and teaching experience. (Doctoral dissertation). Georgetown University.
- Inbar-Lourie, O. (2017). Language assessment literacy. In E. Shohamy, I. Or, & S. May (Eds.), *Language testing and assessment*. Springer.
- Jannati, S. (2015). ELT teachers' language assessment literacy: Perceptions and practices. *The International Journal of Research in Teacher Education*, 6(2), 26–37.
- Kiomrs, R., Abdolmehdi, R., & Naser, R. (2011). On the interaction of test washback and teacher assessment literacy: The case of Iranian EFL secondary school teachers. *English Language Testing*, 4(1), 156–161.
- Koh, K., Burke, L. E. C.-A., Luke, A., Gong, W., & Tan, C. (2017). Developing the assessment literacy of teachers in Chinese language classrooms: A focus on assessment task design. *Language Teaching Research*, 22(3), 264–288.
- Lopez, O. S. (1995). Classroom diversification: An alternative paradigm for research in educational productivity. (Unpublished doctoral dissertation). University of Texas, Austin.
- Mackey, A., Polio, C., & McDonough, K. (2004). The relationship between experience, education, and teachers' use of incidental focus-on-form techniques. *Language Teaching Research*, 8(3), 301–327.
- McNeil, A. (2005). Non-native speaker teachers and awareness of lexical difficulty in pedagogical texts. In E. Llorca (Ed.), *Non-native language teachers: Perceptions, challenges and contributions to the profession* (pp. 107–128). New York, NY: Springer.
- Mede, E., & Atay, D. (2017). English language teachers' assessment literacy: The Turkish context. *Dil Dergisi*, 168(1), 43–60.
- Mertler, C. A. (2003). Preservice versus in-service teachers' assessment literacy: Does classroom experience make a difference? In *Annual meeting of the Mid-Western Educational Research Association*, Columbus, OH.
- Montee, M., Bach, A., Donovan, A., & Thompson, L. (2013). LCTL teachers' assessment knowledge and practices: An exploratory study. *Journal of the National Council of Less Commonly Taught Languages*, 14, 1–32.
- Mok, W. E. (1994). Reflecting on reflections: A case study of experienced and inexperienced ESL teachers. *System*, 22(1), 93–111.
- Muñoz, A. P., Palacio, M., & Escobar, L. (2012). Teachers' beliefs about assessment in an EFL context in Colombia. *Issues in Teachers' Professional Development*, 14(1).

- Newby, D., Allan, R., Fenner, A. B., Jones, B., Komorowska, H., & Soghikyan, K. (2007). *European portfolio for student teachers of languages: A reflection tool for language teacher education*. Strasbourg: Council of Europe.
- O'Loughlin, K. (2013). Developing the assessment literacy of university proficiency test users. *Language Testing*, 30(3), 363–380.
- Ölmezer-Öztürk, E., & Aydın, B. (2018). Investigating language assessment knowledge of EFL teachers. *Hacettepe University Journal of Education*. <https://doi.org/10.16986/HUJE.2018043465>
- Öz, S., & Atay, D. (2015). Turkish EFL instructors' in-class language assessment literacy: Perceptions and practices. *ELT Journal*, 6(1), 25–44.
- Özdemir-Yılmaz, M., & Özkan, Y. (2017). Speaking assessment perceptions and practices of English teachers at tertiary level in the Turkish context. *Language Learning in Higher Education*, 7(2).
- Piasecka, L. (2018). Tinker, tailor...: Creativity in foreign language learning and teaching. In M. Pawlak & A. Mystkowska-Wiertelak (Eds.), *Challenges of second and foreign language education in a globalized world: Second language learning and teaching* (pp. 89–106). Cham: Springer.
- Pill, J., & Harding, L. (2013). Defining the language assessment literacy gap: Evidence from a parliamentary enquiry. *Language Testing*, 30(3), 381–402.
- Popham, J. W. (2011). Assessment literacy overlooked: A teacher educator's confession. *The Teacher Educator*, 46(4), 265–273.
- Mead, N. A., & Rubin, D. L. (1985). *Assessing listening and speaking skills*. Washington, DC: National Inst. of Education (ED).
- Sarıyıldız, G. (2018). A study into language assessment literacy of preservice English as a foreign language teachers in the Turkish context. (Master's thesis). Hacettepe University.
- Shim, K. N. (2009). An investigation into teachers' perceptions of classroom-based assessment of English as a foreign language in Korean primary education. (Unpublished doctoral dissertation). University of Exeter, Exeter.
- Stiggins, R. (1991). Assessment literacy. *Phi Delta Kappan*, 72, 534–539.
- Stiggins, R. J. (1995). Assessment literacy for the 21st century. *Phi Delta Kappan*, 77(3), 238–245.
- Tajeddin, Z., Alemi, M., & Yasaei, H. (2018). Classroom assessment literacy for speaking: Exploring novice and experienced English language teachers' knowledge and practice. *Iranian Journal of Language Teaching Research*, 6(3), 57–77.
- Tamerer, R. (2019). An investigation of Turkish pre-service EFL teachers' language assessment literacy. (Master's thesis). Kocaeli University.
- Vogt, K., & Tsagari, D. (2014). Assessment literacy of foreign language teachers: Findings of a European study. *Language Assessment Quarterly*, 11(4), 374–402.
- Wiggins, J. S., & Pincus, A. L. (1989). Conceptions of personality disorders and dimensions of personality. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 1(4), 305.
- Zaim, M., Refnaldi, & Arsyad, S. (2020). Authentic assessment for speaking skills: Problems and solutions for English secondary school teachers in Indonesia. *International Journal of Instruction*, 13(3), 587–604.