Despite many studies on assessment literacy, scant attention has been devoted to classroom assessment literacy for speaking. To bridge this gap, the present study set out to explore novice and experienced English as a foreign language (EFL) teachers’ classroom assessment literacy for speaking. Twenty-six novice and experienced EFL teachers first completed three scenarios to collect data about their practices for speaking assessment. Afterwards, they went through semi-structured interview to elicit their speaking assessment knowledge. Both scenarios and interviews were framed using Xu and Brown’s (2016) six-component framework of teacher assessment literacy. The focus of this study was on the teacher knowledge base, which is the second component in this framework and consists of knowledge of assessment purposes, content/criteria, and methods. Novice and experienced EFL teachers’ scenario data were content analyzed through a summative approach, which led to the discovery of formative assessment and diagnostic assessment as their knowledge of assessment purposes in practice. Moreover, this analysis shed light on the teachers’ six assessment criteria in practice, including accuracy, fluency, vocabulary, pronunciation, communication, and content. As to the teachers’ assessment methods in practice, the findings revealed their use of role play, monolog, picture story, and questions and answers. The interview data were transcribed and content analyzed through a summative approach as well, which resulted in three main categories of assessment purposes: formative assessment, diagnostic assessment, and summative assessment. Besides knowledge of assessment purposes, the interview data substantiated the scenario findings about the teachers’ assessment criteria and their beliefs about various assessment methods such as role plays, monologs, and authentic tasks. Although divergence between novice and experienced teachers’ beliefs was moderate, the data demonstrated more consistency in the experienced teachers’ assessment literacy for speaking. The findings imply that EFL teachers need awareness of their underlying knowledge about speaking, which is the product of their past education and language teaching experiences.

Keywords: assessment literacy; teacher knowledge; teacher practice; speaking; novice teachers; experienced teachers

© Urmia University Press

ARTICLE HISTORY

Received: 30 Mar. 2018
Accepted: 20 Sept. 2018
Revised version received: 15 June 2018
Available online: 1 Oct. 2018

* Corresponding author: Islamic Azad University – West Tehran Branch, Tehran, Iran
Email address: minoalemi2000@yahoo.com
© Urmia University Press

10.30466 / ijltr.2018.120601
Introduction

Classroom assessment literacy refers to the requisite knowledge and skill for collecting data about learners’ performance along with the ability to apply the assessment outcome to improve learners’ achievement (Campbell, Murphy, & Holt, 2002; Popham, 2005). Classroom assessment literacy has been identified as one of the pre-eminent responsibilities of a teacher (Fulcher, 2012; Leung, 2014; Mertler, 2005). This is due to the fact that one-third to one half of any individual teacher’s time is devoted to classroom assessment (Plake, 1993). In addition, it is a widely held view that classroom assessment literacy is an integral part of teaching profession (Abell & Siegel, 2011; Mertler, 2009) and has a vital role in bringing about improvement in both teachers’ instruction and learners’ level of achievement in learning (Campbell & Collins, 2007; Guskey, 2003; Harris, Irving, & Peterson, 2008). In this regard, Ecclestone, Davis, Mickelson, and Goodburn (2005) argue about the impossibility of good teaching in the absence of good assessment.

On the other hand, the relationship among teachers’ instructional knowledge and beliefs, the implicit and not clearly articulated theories (Rea-Dickins, 2007), and their practices have generated wide interest among scholars for more than two decades. There exists some evidence that teachers’ beliefs have profound impacts on the decisions they make regarding their instruction and, in turn, their practices (Malone, 2013; Pajares, 1992; Squires & Bliss, 2004; Winterbottom, Brindley, Taber, Fisher, Finney, & Riga, 2008). Likewise, Maziarz (2007) maintains that a teacher’s knowledge base and beliefs operate as a filter through which new concepts are processed and interpreted. Moreover, while it has commonly been assumed that there is a complex rather than a unilateral relationship between teachers’ knowledge/beliefs and their practices (Poulson, Avramidis, Fox, Medwekk, & Wray, 2001), investigation of this multiplex system of beliefs and its effect on teachers’ practices is pivotal (Barr, 2001).

One aspect of teachers’ knowledge base is their conceptualization and practice of classroom assessment and its corresponding assessment literacy. The significance of classroom assessment literacy is evident in the number of articles and books published (e.g., Leung, 2014; Malone, 2013; Scarino, 2013; Taylor, 2009; Xu & Brown, 2016). While some researchers have addressed the issues related to teachers’ knowledge (e.g., Brown, 2011; Gebril & Brown, 2013), others have tended to focus on classroom assessment practices (e.g., Calveric, 2010; Ganske, Monroe, & Strickland, 2003; Qian, 2008). However, most of these studies have examined classroom assessment literacy in general or in relation to student learning. Consequently, very few studies have discussed classroom assessment literacy for speaking (Tajeddin, Alemi, & Yousefi, 2017) though speaking performance can be used as evidence of not only speaking ability but also the second language acquisition process (Douglas, 1994). Moreover, few studies have taken into account novice and experienced EFL teachers’ assessment knowledge and practices. Against this backdrop, the main aim of this study was to fill these gaps by exploring novice and experienced EFL teachers’ knowledge and practices of classroom assessment literacy for speaking.

Literature Review

Assessment Literacy

Due to its vital role in language teaching and learning, assessment has always been the center of attention. This is clearly evident in the number of articles published to investigate into language assessment (e.g., Alemi & Khanlarzadeh, 2016; Douglas, 2014, and Sadeghi & Zeinolabedini, 2016, to name a few). Recently, however, literature has witnessed considerable interest in assessment literacy. Consequently, the term assessment literacy has been defined by many scholars. Being the first to define the concept, Stiggins (1995) listed five competencies of sound
assessment: what to assess, why to assess, how to assess, the possible problems of assessment, and how to prevent these problems. Davies (2008), on the other hand, noted skills, knowledge, and principles as the basic components of assessment literacy. Fulcher (2012) also defined assessment literacy as having three interrelated dimensions. In Fulcher's model, while the first dimension refers to the adequate knowledge and skills for designing and evaluating a test, the second one deals with familiarity with the ethical issues. The third dimension of assessment literacy, however, encompasses the capability to place assessment in the wider social, political, cultural frameworks so as to evaluate the consequences of assessment on society and institutions.

Most of the empirical research on assessment literacy began with the seminal document “The Standards for Teacher Competence in Educational Assessment of Students” (AFT, NCME, & NEA, 1990), which encompassed seven competencies. Following its publication, researchers devised instruments to investigate the teachers’ assessment literacy (e.g., Mertler, 2003; Zheng, 2010). For instance, Plake and Impara (1992) used the 1990 standards as a blueprint to devise a new instrument titled Teacher Assessment Literacy Questionnaire (TALQ). The application of this 35-item questionnaire to 555 in-service teachers in the United States revealed their low level of assessment literacy. Similarly, Campbell et al.’s (2002) administration of TALQ to 220 teachers indicated that they were not prepared to assess the learners.

Assessment literacy has been examined in relation to teacher demographics as well. One of the demographic features was language teaching experience, which produced mixed results. While Zhang and Burry-Stock’s (2003) study did not reveal a positive relationship between teachers’ experience and assessment literacy, Mertler’s (2003) study confirmed this relationship. In another study, Mertler (2009) used a parallel mixed method design to analyze the effects of a two-week classroom assessment workshop. The researcher used a questionnaire both as a pre-test and post-test to find out the effects of the workshop which was based on the 1990 Standards. The results of the post-test proved the effectiveness of training and indicated a significant increase in teachers’ assessment literacy.

More recently, researchers have called for an updated version of the assessment literacy to keep up with latest developments in formative assessment such as self-assessment and peer-assessment (e.g., Brookhart, 2011; Stiggins, 2010). Xu and Brown (2016), for instance, synthesized and analyzed 100 studies in assessment literacy and proposed a framework named teacher assessment literacy in practice (TALiP). The framework encompasses six components: the knowledge base, teacher conceptions of assessment, institutional and socio-cultural contexts, teacher assessment literacy in practice (TALiP), teacher learning, and teacher identity. Having categorized TALiP as the core component of this conceptualized framework of assessment literacy, they argued that all the six components are interrelated and the model has a cyclical nature.

Research on Teachers’ Assessment Knowledge and Practice

While literature on teachers’ assessment knowledge and practice regarding the skill of speaking is not rich, a considerable amount of literature has been published on the relationship between teachers’ knowledge and practices and their demographic characteristics (e.g., Calveric, 2010; Davison, 2004). In 1996, Cizek, Fitzgerald, and Rachor conducted an investigation into 143 elementary and secondary school teachers to explore if there is any significant relationship between their assessment practices and their gender and years of experience. The findings indicated that teachers’ assessment practices could not be predicted from their gender and years of experience.
Some studies have been conducted to investigate the relationship between teachers’ assessment knowledge and assessment practices regarding learning and teaching (e.g., Black & William, 1998; Kahn, 2001; Popham, 2009). Kahn (2001), for instance, examined teacher-centered assessment materials to explore teachers’ knowledge about teaching and learning. Careful analysis of both teachers’ assessment materials and their provided comments uncovered the methodology behind the assessment materials. While some assessment materials made use of constructive methodology, the rest still utilizing traditional ones. Kahn came to the conclusion that teachers’ assessment practices were under the influence of their knowledge about what constitutes learning.

As the preceding literature shows, a few studies have been conducted to investigate different aspects of assessment literacy. Nevertheless, these attempts have been restricted to the investigation of the components of assessment literacy and the relationship between teachers’ assessment literacy and their demographic features. Moreover, few studies have taken into account novice and experienced EFL teachers’ knowledge and practices of speaking assessment. To bridge this gap, this study sought to broaden the current understanding of assessment literacy by investigating novice and experienced EFL teachers’ assessment knowledge and practice regarding speaking. To achieve this aim, the following questions were raised:

1. What knowledge do novice and experienced EFL teachers’ have about assessing speaking in terms of purposes, content, and methods?
2. What are novice and experienced EFL teachers’ scenario-based practices of speaking assessment in terms of purposes, content, and methods?

Method

Participants

A total of 13 novice and 13 experienced EFL teachers participated in this study on a voluntary basis. While the experienced EFL teachers had been teaching English in language institutes for over ten years, the novices had up to three years of teaching experience in language institutes. Initially, emails were sent to 40 novice and 40 experienced EFL teachers to invite them to participate in the study. Subsequently, appointments were made with each of those who accepted the invitation. The participants first completed the scenarios developed by the researchers and then participated in a face-to-face interview conducted by the third author. The participants’ age ranged from 20 to 44 and all were female except one. The reason for the female majority was that all the male teachers except one rejected the invitation although both female and male teachers were requested to participate in the study. Another reason was that most of the teachers at the institute were female. Therefore, it was decided to base the work on the teachers’ accessibility and willingness to participate.

Instrumentation

The instruments used in the present study, which is part of a larger study, were scenarios and semi-structured interviews. The scenarios were administered before the interviews so that the participants could authentically engage in assessment practice before verbalizing their knowledge in the interviews.
Scenarios

Scenarios were used to collect data about the participants’ practices for speaking. The scenarios employed in this study were developed by the researchers themselves based on the second components of the first dimension of Xu and Brown’s (2016) framework of teacher assessment literacy. This component consists of knowledge of assessment purposes, content/criteria, and methods. Informed by this component, three different scenarios were prepared, each of which was intended to elicit a single aspect of classroom assessment literacy. It should be noted that the researchers met two sessions to prepare the three scenarios. The scenarios regarding knowledge of assessment purposes, content, and methods are given below, respectively:

Scenario #1: You have a class with elementary learners in a language institute and one day the supervisor comes to the class for an observation. At the end of the class, he invites you to the office and asks you about the purposes you have in speaking assessment. How do you describe them?

Scenario #2: Consider you have a class with elementary learners in a language school. And your language learners are supposed to give a summary of a story they had in their text book. You ask one of them to give the summary. What aspects of her performance do you assess?

Scenario #3: It is the first session of your class with a group of elementary language learners in a language institute and you don’t have enough information about their speaking skills. What tasks do you use to assess their speaking?

Interviews

Semi-structured interviews were conducted in the present study since, according to Dörnyei (2007), a semi-structured interview is a “compromise” where the researcher knows the key questions prior to the interview and at the same time it allows interviewees to assert their own views and experiences. Therefore, an interview guide generated from Xu and Brown’s (2016) model of language assessment literacy was used to uncover the participants’ knowledge of classroom assessment of speaking. The second component of this model, i.e. knowledge of assessment purposes, content, and methods, was the main focus of the present study. The researchers met two sessions to decide about the kinds of questions to be embedded in the interview. Eventually, they decided to prepare three questions for assessment purposes, content, and methods as given below:

Interview question 1: How do you describe the purpose of speaking assessment in your class?

Interview question 2: What do you assess mostly in your learners’ speaking performances?

Interview question 3: What tasks do you use to assess your learners’ speaking?

All the interviews were conducted in the participants’ native language, Persian. Each interview lasted 10 to 20 minutes; but as this study is part of a larger study, only 5 to 10 minutes of each interview were related to the present study. Sufficient elaboration and examples were provided to explain the parts where misunderstandings occurred. In the end, the recorded interviews were transcribed for further reflection and analysis.
Data Collection and Analysis

After 26 out of 80 invitations for participation in this study were accepted, the participants were invited to the third author’s office for both scenario completion and semi-structured interviews. Each of the teachers first completed the scenarios and afterwards went through the semi-structured interview based on the prepared scenarios and the interview guide. During scenario completion and interviews, more elaboration was provided in case misunderstanding happened. The interviews were audio recorded and subsequently transcribed. The final transcription was e-mailed to the participants to be checked by them as well. They were asked to read the transcriptions carefully to either add or delete the parts which needed modification. They all confirmed the content of the transcriptions and no change was made accordingly.

To answer the two research questions, the qualitative procedure of summative content analysis was employed to discover the major categories in the data for each of the research questions. To this end, after data collection, the researchers coded the data. The codes were derived by the third author who started qualitative data analysis using paper and pencil on hard copies of the data. For instance, while attempting to explore the main categories for the first research question regarding EFL teachers’ knowledge of speaking assessment purposes, she found out relationship among the phrases the participants used such as “every moment,” “all the time,” or “to help them learn better.” She decided to group them under the main category of “formative assessment” as they reflected online assessment. Afterwards, the three authors met to reflect on the derived codes and categories from different perspectives. The only modification they made was in the section of teachers’ knowledge of assessment methods where “picture telling” was changed to “picture story”. Moreover, the strategy of member checking proposed by Birt, Scott, Cavers, Campell, and Walter (2016) was used in some parts of the scenario phase in order to ensure that the researchers’ understanding of the answers to the scenarios was confirmed by the participants themselves. Two experts in the field of applied linguistics were subsequently asked to judge and comment about the final version of the extracted main categories and revisions were applied where needed.

Results

Beliefs about Speaking Assessment

The first research question was aimed at discovering novice and experienced EFL teachers’ knowledge of speaking assessment. To address the question, all the participants were interviewed individually. The gathered data were content analyzed carefully by the researchers to derive the main themes. As indicated previously, since each research question encompassed the three variables of knowledge of assessment purposes, content, and methods, a separate interview question was prepared for each of these components. Therefore, to gain a richer insight into the teachers’ knowledge of each variable, each interview questions was analyzed separately.

The first interview question was: How do you describe the purpose of speaking assessment in your class? It was focused on novice and experienced EFL teachers’ knowledge of speaking assessment purposes. The content analysis of this part resulted in three main categories of assessment purposes: formative assessment, diagnostic assessment, and summative assessment. Table 1 describes these three assessment purposes.
Table 1

Novice and Experienced EFL Teachers’ Knowledge of Speaking Assessment Purposes

<table>
<thead>
<tr>
<th>Main Category</th>
<th>Experienced EFL Teachers</th>
<th>Novice EFL Teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Frequency</td>
<td>Frequency in Total</td>
</tr>
<tr>
<td>Formative Assessment</td>
<td>13</td>
<td>100%</td>
</tr>
<tr>
<td>Diagnostic Purposes</td>
<td>7</td>
<td>54%</td>
</tr>
<tr>
<td>Summative Assessment</td>
<td>3</td>
<td>23%</td>
</tr>
</tbody>
</table>

**Formative Assessment.** This kind of assessment purpose was specified by the majority of both experienced teachers (ET) and novice teachers (NT) and they all had a high opinion of it. All the experienced teachers (100%) and most of the novice ones (92%) appreciated this kind of assessment purpose. For instance, an experienced teacher (T2) described her assessment during the learners’ learning process this way:

> Every moment I assess their speaking… when I’m teaching or when they are doing a role play or so… I monitor them to see if they need any help or correction or more explanation … this way I can help them improve and learn better. [ET2]

As it is evident from the above excerpt, this teacher used formative assessment to help learners improve continuously.

**Diagnostic Assessment.** The second assessment purpose was undertaken to assess learners’ performance and identify their weaknesses and strengths in order to inform, plan, or modify the next steps in teaching. While 7 out of 13 experienced teachers (54%) argued for diagnostic assessment as their speaking assessment purpose, only 2 out of 13 novice ones (15%) noted it. For instance, an experienced teacher (ET1) elaborated on her diagnostic assessment purpose and commented:

> Every session … before I start teaching… I asked them some questions … for example I want them to talk about their home and while they are speaking… I check to see what are their weak points and strong points… I mean this way I gain a better understanding of what they know and I decide how to teach that day. [ET1]

An example of how novice teachers applied diagnostic assessment can be found in the statement by NT12.

> I assess them to find their weak points and strong points so that I can help them improve little by little. This also helps me to make a decision about the things I want to teach next sessions. [NT12]

In both responses, the teachers favored diagnostic assessment, but what made a difference between the two was the fact that five of the experienced teachers who reported on diagnostic assessment made use of more professional terms while the novice one talked about her experience in simple words. Another finding was that experienced teachers tended to utilize the result of their diagnostic assessment as a base to manage the same session while novice teachers would make use of them in the next sessions. This might be related to the ability of experienced teachers to make better decisions.
Summative Assessment. This assessment purpose happened at the end of the term and sought to monitor educational outcomes. Three experienced teachers (23%) and two novice ones (15%) referred to the application of summative assessment as their speaking assessment purpose. An example derived from one of the experienced EFL teachers’ responses is given below.

At the end of the term, each of them has to take a speaking test which lasts about five to ten minutes. Some questions are prepared before and I ask each of them those questions. [ET6]

A novice teacher also reported the use of summative assessment as she commented:

I assess their speaking at the end of the term too. [NT3]

As these two statements show, the experienced teacher knew not only about the importance of summative assessment but also about the way to implement it. The novice teacher simply reported on her use of summative assessment with no mention of the ways she could do it.

The second interview question was: What do you assess mostly in your learners’ speaking performances? It was aimed at discovering the participants’ knowledge of assessment content. The analysis of interview data led to the discovery of eight main criteria for speaking assessment discussed by the participants, as shown in Table 2.

Table 2
Novice and Experienced EFL Teachers’ Knowledge of Speaking Assessment Criteria

<table>
<thead>
<tr>
<th>Main Category</th>
<th>Experienced EFL Teachers</th>
<th>Novice EFL Teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Frequency</td>
<td>Percentage</td>
</tr>
<tr>
<td>Accuracy</td>
<td>9</td>
<td>69%</td>
</tr>
<tr>
<td>Fluency</td>
<td>10</td>
<td>77%</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>7</td>
<td>54%</td>
</tr>
<tr>
<td>Pronunciation</td>
<td>7</td>
<td>54%</td>
</tr>
<tr>
<td>Communication</td>
<td>6</td>
<td>46%</td>
</tr>
<tr>
<td>Interaction</td>
<td>3</td>
<td>23%</td>
</tr>
<tr>
<td>Pragmatics</td>
<td>4</td>
<td>31%</td>
</tr>
<tr>
<td>Content</td>
<td>3</td>
<td>23%</td>
</tr>
</tbody>
</table>

Accuracy. Both novice and experienced teachers paid almost equal attention to the importance of their learners’ accuracy. Nine out of 13 experienced teachers (69%) and 10 out of 13 novice ones (77%) valued accuracy in their learners’ speaking assessment. However, what differed was the way they approached the issue. As the following excerpts show, experienced teachers employed certain strategies while dealing with learners’ accuracy. One of the experienced EFL teachers uttered:

When a learner makes a grammatical error, I don’t say them it is grammatically wrong…what I do is… I actually ask the other learners to correct him and if they can’t do it, I do it myself. [ET12]

Most novice teachers only announced that they assessed accuracy, but they failed to explain how they did it. This is obvious in an excerpt from NT11:

I assess both their accuracy and their pronunciation. (NT11)
Fluency. Whereas 10 experienced teachers (77%) highlighted fluency, only 5 novice teachers (39%) discussed it. A novice teacher voiced her concern like this:

I prefer my learners speak more fluently than more accurately. [NT12]

Vocabulary. Similar to the previous criteria, experienced teachers pointed to vocabulary more than novice teachers. While 7 experienced teachers (54%) named vocabulary as what they assessed mostly, 2 novice ones (15%) did so. According to one of the experienced teachers:

I also check if they use the new vocabulary they’ve learnt while speaking about a topic. If they are talking about a topic … seasons for example… I check if they use those new words or not. [ET7]

Pronunciation. The fourth criterion was pronunciation to which both novice and experienced teachers paid the same amount of attention (54%). In this regard, a novice teacher stated:

I assess the learners’ pronunciation in all their performances and correct their mispronunciations as soon as they happen. [NT6]

Communication. The fifth derived criterion, i.e. communication, addressed the creativity learners can have to make themselves understood while communicating with their classmates. About 46% of the experienced teachers (6 out of 13) named it as a factor in speaking assessment. However, none of the novice teachers uttered a single word about it. According to an experienced teacher:

You know, I was not conscious of communication before. But when my husband talked about his experiences of living abroad, I changed my method… now I teach my learners to be careful and see whether their classmates understand them or not. This is a very important issue in my assessment. So in role plays, I clearly ask them to be more careful about this aspect. [ET13]

Interaction. This criterion refers to the learners’ ability to interact with others effectively, namely understanding the interlocutor’s question and being able to answer that question appropriately. This criterion was within the experienced teachers’ knowledge of assessment content. Three experienced teachers (23%) valued this assessment content, as one of them asserted:

In role plays, for example, I want them to know how a conversation goes on… how it works. I don’t let one of them just asks questions and the other one just gives him an answer. Both of them should be involved in asking and answering the questions. They should learn the arts of listening, responding, and speaking. [ET2]

Pragmatics. Like the previous two criteria, i.e. communication and interaction, pragmatics was specified only by the experienced EFL teachers. That is, four experienced teachers (31%) referred to it as their speaking assessment criterion. According to ET3:

If there is a situation in their role play and pragmatics can be checked and I have taught them before … I will assess that too… things like initiations and turn taking. [ET3]

Content. Both experienced (23%) and novice (8%) teachers referred to this criterion. The only novice teacher [NT13] who mentioned her use of this criterion described her focus on it this way:

“… the learners should not speak off the topic.”
The third interview was raised to explore novice and experienced teachers’ knowledge of assessment methods. Summative content analysis of the interviews uncovered eight assessment methods which are presented in Table 3.

Table 3
Novice and Experienced EFL Teachers’ Knowledge of Speaking Assessment Methods

<table>
<thead>
<tr>
<th>Main Category</th>
<th>Experienced EFL Teachers</th>
<th>Novice EFL Teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Frequency</td>
<td>Percentage</td>
</tr>
<tr>
<td>Role Play</td>
<td>13</td>
<td>100%</td>
</tr>
<tr>
<td>Summary</td>
<td>7</td>
<td>54%</td>
</tr>
<tr>
<td>Questions and Answers</td>
<td>3</td>
<td>23%</td>
</tr>
<tr>
<td>Monolog</td>
<td>2</td>
<td>15%</td>
</tr>
<tr>
<td>Authentic Task</td>
<td>2</td>
<td>14%</td>
</tr>
<tr>
<td>Problem Solving</td>
<td>1</td>
<td>8%</td>
</tr>
<tr>
<td>Picture Story</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td>Game</td>
<td>0</td>
<td>0%</td>
</tr>
</tbody>
</table>

**Role Play.** While all the experienced teachers (100%) included this category in their assessment methods, only seven novice teachers (54%) gave weight to it. A novice teacher commented: “I usually use role plays… they are good methods I think” [NT6].

**Summary.** The second method involved learners in giving a summary of what they had read or listened to. Seven out of 13 experienced teachers (54%) and 5 out of 13 novice ones (38%) made use of this assessment method to assess their learners’ speaking. However, there was a difference in their approaches to summary. Whereas five out of seven experienced teachers used oral prompts like short clips, cartoons, and conversations, five novice teachers used written prompts. An experienced teacher, for example, stated, “They listen to a conversation and then I ask them to give a summary of that conversation”[ET4].

**Questions and Answers.** Both experienced and novice teachers had almost an equal awareness of this assessment method in which the teacher asks each learner some questions related to the lesson or some general questions to assess their speaking. Three experienced teachers (23%) and two novice ones (15%) talked about this assessment method. For instance, a novice teacher maintained, “I sometimes ask them some questions related to the topic of that unit” [NT3].

**Monolog.** This assessment method focused on spoken performances where a learner gave a talk in front of the other learners. Novice teachers made use of this method twice as many times as experienced ones did. That is, two experienced teachers (15%) and four novice ones (30%) argued for the use of this speaking assessment method. As a novice teacher reported, “I ask one of the learners to come to the board and talk about his last vacation for instance” [NT7].

**Authentic Task.** Five out of 13 experienced teachers (38%) supported the use of authentic tasks while novices gave no weight to this method. According to an experienced teacher:

> Most of the learners want to migrate or they travel abroad a lot … I ask them to imagine a situation like ordering in a restaurant then we work on that… this is one of the methods I use to assess them. [ET5]
Problem Solving. This assessment method was used only by one of the experienced teachers. This is reflected in the statement below:

> It depends on the topic of that session or lesson, but I make tasks myself. For example, if their lesson is related to house and rooms… I group them and ask them to imagine they buy a new house and they want to buy some furniture with a certain amount of money, or they want to decorate the house… they have to discuss and convince each other to put for example the sofa under the window or not. Then they come to the board and perform it. The best performance will win… so this way a kind of competition happens too.

[ET13]

Picture Story. This assessment purpose was used by six novice teachers (46%) only. As one of them voiced, “I use the pictures in their books, or I take some pictures to class and ask them to speak about them” [NT10].

Game. The last assessment purpose was specified by only one of the novice teachers as she stated, “We play games sometimes if the students like it of course” [NT2]. She went further and reported that she usually paired her language learners and asked each group to come to the board. Then, she gave one of them a picture to describe it to her partner. The partner’s responsibility was to draw it on the board.

Novice and Experienced EFL Teachers’ Scenario-Based Practices

The aim of the second research question was to explore novice and experienced teachers’ scenario-based practices of speaking assessment. Similar to the first research question, this research question was examined regarding the three components of speaking assessment literacy. To this end, teachers’ answers to each of the three scenarios were content analyzed through a summative approach. In order to obtain a revealing picture of the results in this section, scenarios are analyzed one by one.

Scenario 1: You have a class with elementary learners in a language institute and one day the supervisor comes to the class for an observation. At the end of the class, he invites you to the office and asks you about the purposes you have in speaking assessment. How do you describe them?

This scenario set out to explore novice and experienced teachers’ scenario-based practices of assessment purposes. Content analysis through a summative approach resulted in two themes: formative assessment and diagnostic assessment.

Formative Assessment. All the 13 experienced teachers (100%) and 12 novice teachers (92%) wrote about their online and ongoing speaking assessment. As one of the experienced EFL stated:

> I will tell him that my purpose is to assess them every moment they are speaking and this way I can help them improve easier… and prevent fossilization. [ET10]

A novice EFL teacher also asserted:

> I always assess their speaking as they are performing tasks or speaking with me. If I don’t act this way, they won’t make any progress. [NT12]
**Diagnostic Purposes.** While 8 out of 13 experienced teachers (61%) tended to draw on diagnostic assessment, only one novice teacher (8%) practiced it. An experienced teacher’s answer to this scenario was:

> I would say that speaking is the product of whatever I try to do in my class. Assessment is the last process but not the least. I can check all their weaknesses and strong points through speaking assessment. [ET1]

A novice teacher, on the other hand, jotted down:

> To help them find out their errors and weak points. [NT6]

A summary of the above themes is presented in Table 4.

<table>
<thead>
<tr>
<th>Main Category</th>
<th>Experienced EFL Teachers</th>
<th>Novice EFL Teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Frequency</td>
<td>Percentage</td>
</tr>
<tr>
<td>Formative Assessment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diagnostic Purposes</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 4

*Novice and Experienced EFL Teachers’ Scenario-Based Practices as to Speaking Assessment Purposes*

The second scenario sought to uncover teachers’ practices about assessment criteria. To do that, all the gathered data were analyzed through a summative approach, which led to the discovery of six assessment criteria: accuracy, fluency, vocabulary, pronunciation, communication, and content.

**Fluency.** This assessment criterion was specified by 10 experienced teachers (77%) and 3 novice ones (23%). For example, an experienced teacher commented, “For summary, both fluency and accuracy are important” [ET13]. In a similar vein, a novice teacher answered, “How fluent he/she is” [NT9].

**Accuracy.** The second criterion was accuracy to which both the experienced and novice teachers attended equally (62%). According to an experienced teacher, “The grammar they use while speaking, whether it is rich or not” [T8]. A novice just asserted, “Grammatical structures” [NT12].

**Vocabulary.** The third criterion was vocabulary, which attracted more attention from novice teachers. Whereas seven novice teachers (54%) noted their use of vocabulary, experienced ones gave comparatively less weight to it (five of them: 38%). This is how a novice EFL teacher replied, “The vocabulary he uses” [NT11].

**Pronunciation.** Pronunciation was highlighted by five experienced teachers (38%) and two novice ones (15%). An experienced EFL teacher, for instance, wrote, “The ability to produce the right sounds and pronunciation” [ET6].
Communication. This criterion was specified by four experienced teachers (31%) and two novice ones (15%). As an experienced teacher answered, “His eye contact with the other classmates” [ET10].

Content. The last extracted criterion was content, which nine novice teachers (69%) and four experienced ones (31%) favored in their teaching practices. In this regard, a novice EFL teacher maintained, “Just understanding the main points are enough” [NT8].

Table 5 depicts teachers’ scenario-based practices about assessment content.

Table 5
Novice and Experienced EFL Teachers’ Scenario-Based Practices as to Speaking Assessment Content/Criteria

<table>
<thead>
<tr>
<th>Main Category</th>
<th>Experienced EFL Teachers</th>
<th>Novice EFL Teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluency</td>
<td>10 77%</td>
<td>3 23%</td>
</tr>
<tr>
<td>Accuracy</td>
<td>8 62%</td>
<td>8 62%</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>5 38%</td>
<td>7 54%</td>
</tr>
<tr>
<td>Pronunciation</td>
<td>5 38%</td>
<td>2 15%</td>
</tr>
<tr>
<td>Communication</td>
<td>4 31%</td>
<td>2 15%</td>
</tr>
<tr>
<td>Content</td>
<td>4 31%</td>
<td>9 69%</td>
</tr>
</tbody>
</table>

Scenario 3: It is the first session of your class with a group of elementary language learners in a language institute and you don’t have enough information about their speaking skills. What tasks do you use to assess their speaking?

The third scenario was aimed at exploring novice and experienced teachers’ practices in using assessment methods. The teachers’ answers to this scenario, content analyzed through a summative approach, yielded four major themes: role play, monolog, picture story, and questions and answers.

Role Play. The first derived method was specified by 11 experienced teachers (85%) and only one novice teacher (8%). As an experienced teacher wrote, “I ask them to do a role play” [ET6].

Monolog. Both experienced and novice teachers equally accentuated monolog (69%). A novice teacher noted, “To talk about themselves” [NT11].

Questions and Answers. The third main method, question and answer, was specified by three experienced teachers (23%) and one novice teacher (8%). As an experienced teacher asserted, “… or I ask them some questions” [ET4].

Picture Story. As the last method, picture story was used by two novice teachers (15%). One of them wrote, “Bring some pictures to class and ask them to talk about those pictures and make a story about it” [NT2].
A summary of novice and experienced teachers’ scenario-based practices in using assessment methods is given in Table 6.

Table 6
Novice and Experienced EFL Teachers’ Scenario-Based Practices about Speaking Assessment Methods

<table>
<thead>
<tr>
<th>Main Category</th>
<th>Experienced EFL Teachers</th>
<th>Novice EFL Teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Frequency</td>
<td>Percentage</td>
</tr>
<tr>
<td>Role Play</td>
<td>11</td>
<td>85%</td>
</tr>
<tr>
<td>Monolog</td>
<td>9</td>
<td>69%</td>
</tr>
<tr>
<td>Questions and Answers</td>
<td>3</td>
<td>23%</td>
</tr>
<tr>
<td>Picture Story</td>
<td>0</td>
<td>0%</td>
</tr>
</tbody>
</table>

Discussion

Many studies have shed light on teachers’ assessment literacy (e.g., Brookhart, 2011; Stiggins, 2010). However, studies which concentrated on assessment literacy for speaking are rather scarce (e.g., Tajeddin et al., 2017). Moreover, to the best of our knowledge, no study has taken into account novice and experienced EFL teachers’ knowledge and practices of their assessment literacy for speaking. Hence, this study was conducted to investigate novice and experienced EFL teachers’ knowledge and practices of speaking assessment using Xu and Brown’s (2016) framework of assessment literacy with a focus on the knowledge of assessment purposes, criteria, and method, which constitute a component of teacher knowledge base.

The first objective of this study was to discover novice and experienced teachers’ knowledge of speaking assessment regarding assessment purposes, criteria, and methods. As to the knowledge of assessment purposes, the results of this study showed that both novice and experienced teachers’ knowledge was centered on formative assessment, diagnostic assessment, and summative assessment. These findings are in line with Kahl, Hofman, and Bryant’s (2013) study in which they identified formative and classroom summative as two varieties of assessment purposes that assessment literate teachers must know. Furthermore, the findings are supported by Xu and Brown’s (2016) study in which they referred to formative and summative assessment as two assessment purposes.

Besides knowledge of assessment purposes, this study documented teachers’ knowledge of assessment criteria. The analysis of the data led to the discovery of eight main categories of assessment criteria in the teachers’ assessment literacy. Both novice and experienced teachers pointed to accuracy, fluency, vocabulary, pronunciation, and content. Moreover, experienced teachers highlighted communication, interaction, and pragmatics. Further analysis of the data revealed that although both novice and experienced teachers specified five common assessment criteria, experienced teachers demonstrated a better knowledge of the criteria they used. Novice teachers, on the other hand, tended to manifest a poor awareness of the criteria. This finding is similar to the results of Okas’ (2016) study where experienced and novice teachers’ practical knowledge regarding delivery of lessons was analyzed. The findings revealed that although both groups of teachers referred to the same categories, experienced teachers further specified the strategies they used.
The current study provided evidence on EFL teachers’ knowledge of eight speaking assessment methods. Whereas role play, summary, questions and answers, and monolog were acknowledged by both experienced and novice teachers, authentic task and problem solving were utilized by experienced teachers, and picture story and game were specified by novice teachers only. These findings resonate with Brown’s (2004) taxonomy of speaking tasks with five categories: imitative, intensive, responsive, interactive, and extensive. Role play and questions and answers fall within responsive tasks in Brown’s classification. Monolog and summary, on the other hand, exemplify extensive tasks in his taxonomy of speaking tasks. In addition, the findings are supported by a study conducted by Lee (2010) where EFL teachers’ knowledge regarding speaking assessment methods turned out to be responsive, extensive, interactive, and intensive. While responsive and extensive tasks have equivalents in the present study, interactive and intensive tasks have not been identified in our study. It should be noted that Brown’s (2004) imitative tasks were absent in both Lee’s (2010) and our study.

The second objective of this study was to investigate EFL teachers’ practices of speaking assessment regarding assessment purposes, criteria, and methods. As to their knowledge of assessment purposes in practice, the main categories derived from the teachers’ scenario responses included formative assessment and diagnostic purposes. The findings of this part are in agreement with the Cross Sectoral Assessment Working Party (2011) which set out to support Australian teachers in developing functional assessment practices. They listed formative and summative assessment as assessment purposes. They also placed diagnostic assessment inside formative one. Consequently, it can be argued that Cross Sectoral Assessment Working Party’s classification of assessment purposes is somehow similar to the finding of this study. The difference is that while it placed diagnostic assessment inside formative assessment, diagnostic assessment stood alone in the present study and summative assessment was not found in this study.

For their knowledge of assessment criteria, both novice and experienced teachers highlighted fluency, accuracy, vocabulary, pronunciation, communication, and content. However, their stated knowledge of interaction and pragmatics features of speaking assessment was absent in their scenario-based practices. The reason could lie in the fact that the scenario whose goal was to elicit their practice about assessment criteria was focused on a situation in which the teachers could not think about the assessment criteria like interaction or pragmatics. For instance, the second scenario was about giving the summary of a story and hence it was difficult to think of interaction or pragmatics as an aspect of their learners’ performance. Consequently, similar to Flores, López, Gallegos, and Barojas (2000) and Pajares (1992), we can argue that it was the contextual features which made a gap between teachers’ knowledge and practices.

Overall, while the matches and mismatches between experienced and novice EFL teachers were moderate, a closer look into the findings demonstrates more consistency in experienced teachers’ assessment knowledge and practice than in novice teachers’. Considering experienced teachers’ knowledge and practice regarding assessment purposes, we would realize that they referred to formative assessment and diagnostic assessment as their speaking assessment purposes in both their practice and their knowledge. Despite this convergence between their knowledge and practices, the only inconsistency was the difference in the frequency of their use. That is, the frequency with which they referred to formative assessment in their stated knowledge and scenario-based practices amounted to 13, but the difference between their knowledge and practices in diagnostic assessment was 1. Regarding novice teachers’ knowledge about speaking assessment purposes, they had three assessment purposes in their knowledge base, including formative, diagnostic, and summative assessment, while they referred to the two purposes of formative and diagnostic assessment in their scenario-based practice. In addition, although they expressed the same attention to formative assessment in both their knowledge and practices, great
inconsistency was observed in the frequency of summative assessment as it was completely absent in their practice. The reason might be that novice teachers are usually under pressure from the supervisors and other administrators to teacher in a pre-determined way although it might different from their implicit or explicit knowledge. It follows that the context where teachers work can play a role. If they feel less imposition from the stakeholders and are more autonomous, they may act in accordance with their knowledge base. This argument has also been supported by Tsagari (2011) and Turner and Purpura (2016), who discussed the importance of local and external contextual elements while examining teachers’ knowledge and practices. Experienced teachers, on the contrary, may teach different proficiency levels, which helps them reflect on their actions in various contexts. These reflections along with more teacher autonomy might help them shape a more consistent system of knowledge and practices over time. Although the focus of this study was speaking assessment literacy, its findings also support Mertler’s (2003) research which uncovered the positive relation between teachers’ experience and assessment literacy.

Conclusion

This study shed light on the novice and experienced EFL teachers’ knowledge and practices of speaking assessment purposes, criteria, and methods. The summative content analysis resulted in three main categories of assessment purposes: formative assessment, diagnostic assessment, and summative assessment. Besides shedding light on assessment purposes, this study provided evidence on the teachers’ knowledge of assessment content, including accuracy, fluency, vocabulary, pronunciation, communication, interaction, pragmatics, and content. Regarding the novice and experienced teachers' knowledge of assessment methods, the findings showed their knowledge of various assessment methods such as role plays, monologs, and authentic tasks. From these findings, it can be concluded that novice and experienced teachers have both shared and divergent speaking assessment literacy.

The implication of this study is for both teacher educators and teachers. Teacher educators should take into account teachers’ knowledge base about assessment and their assessment literacy. As Cabaroglu and Roberts (2000) contend, teachers’ knowledge and beliefs have effects on their practices. If the teacher education programs allow teachers to freely discuss their knowledge and beliefs about assessment, teacher educators can find the opportunities to make teachers aware of their knowledge base and the probable conflicts among their knowledge and practices. Teacher educators should also highlight the link between teaching and assessment and the importance of assessment in learners’ success by creating situations where teachers’ awareness of assessment in general and assessment literacy of the four skills can be enhanced. Also, teachers should be more reflective on their assessment knowledge and practices and try to bridge the gap between the two.

This study like other studies had a number of limitations. The first limitation was the number of participants. More informative results about teacher assessment literacy will be gained if more participants are engage. The second limitation came from the female majority of the participants in this study. Different results may emerge if other studies involve rather equal samples of male and female teachers. The last limitation was our focus on one subcomponent of Xu and Brown’s (2016) framework of assessment literacy. A better view of teachers’ assessment literacy for speaking will be gained if all the components are examined. The findings from this study suggest some lines of research on novice and experienced English language teachers’ classroom assessment literacy for other skills, including listening, reading, and writing. Moreover, English language teachers’ classroom assessment literacy can be investigated based on their gender and educational background.
References


Zia Tajeddin is Professor of Applied Linguistics at Tarbiat Modares University, Iran, where he teaches doctoral courses in Discourse Analysis, Interlanguage Pragmatics, and Language, Culture, and Identity. He is the editor of Applied Pragmatics (John Benjamins) and Journal of Second Language Teacher Education, and sits on the editorial/review boards of journals such as RELC Journal and TESL-EJ. His research interests center on L2 pragmatic instruction and assessment, classroom discourse analysis, teacher identity and cognition, and EIL/ELF. He has presented papers in international conferences and published papers in many journals, including The Language Learning Journal, Journal of Intercultural Communication Research, RELC Journal, Australian Journal of Teacher Education, Language and Intercultural Communication, Journal of Language, Identity & Education, The Asia-Pacific Education Researcher, TESL-EJ, and TESL Canada Journal.

Minoo Alemi is Assistant Professor of Applied Linguistics at Islamic Azad University, West Tehran Branch, Iran, and a post-doctoral associate at Sharif University of Technology (SUT), Iran. She is the founder of Robot-Assisted Language Learning (RALL) and the co-founder of Social Robotics in Iran. She is the associate editor of Applied Pragmatics (John Benjamins) and is on the editorial/review boards of many journals, including British Journal of Educational Technology, BRAIN, LIBRI, and Scientia Iranica. Her areas of interest include discourse analysis, interlanguage pragmatics, materials development, and robot-assisted language education. She has presented papers in many international conferences and published papers in journals such as Journal of Intercultural Communication Research, The Asia-Pacific Education Researcher, TESL Canada Journal, and International Journal of Social Robotics.

Hasti Yasaei is a Ph.D. candidate of TEFL at Islamic Azad University, West Tehran Branch, Iran. She has been teaching English at university for a few years. Her areas of interest include teaching methodology, second language acquisition, teacher education, and pragmatics. She has presented papers in national and international conferences such as TELLSI and IELTI.