EFL Teachers' Assessment Literacy of Young Learners: Findings
from a Small-Scale Study[*]

İngilizce Öğretmenlerinin Çocuklara Yönelik Değerlendirme Okuryazarlığı: Küçük Ölçekli Bir Çalışmanın Bulguları

Sabahattin YEŞİLÇINAR[**] iD   Galip KARTAL[***] iD

ABSTRACT: Assessing language skills is crucial for checking the quality of teaching. Teacher assessment literacy plays a vital role in deciding on the quality of classroom assessment. Although assessment is a time-taking process, the assessment of foreign language speaking and writing with their unique characteristics requires more time, effort, and expertise. Moreover, assessing the language production of young learners requires much more attention because the assessment is a part of teaching and has cognitive and socio-cultural foundations. Therefore, this small-scale qualitative research investigated three English teacher's views and practices on how they assess their young learners' speaking and writing skills. Data were collected via semi-structured interviews and subjected to content analysis. The findings revealed the mismatches between teachers' views and practices. It was found that the teachers did not adopt the assessor identity due to various barriers; namely, problems in teacher training and development, avoiding assessing productive skills, and the necessity for rater training. The findings suggest that the assessment literacy of the teachers needs to be improved. Therefore, pre-service teacher education programs should emphasise the identity development of language teachers as assessors. Additionally, in-service training is required for enhancing the assessment literacy of the language teachers.

Keywords: Assessment literacy, language assessment, young learners, assessing writing, assessing speaking.

ÖZ: Dil becerilerinin değerlendirilmesi, öğretimin kalitesini kontrol etmek için çok önemlidir. Benzer şekilde, öğretmen değerlendirme okuryazarlığı sınıf değerlendirmelerinin kalitesine karar vermede hayati bir rol oynamaktadır. Değerlendirme zaman alıcı bir süreç olmasının yanı sıra, yabancı dilde konuşma ve yazmanın kendine has özellikleri ile değerlendirilmesi daha fazla zaman, çaba ve uzmanlık gerektirir. Ayrıca, çocukların dil üretimini değerlendirmek çok daha fazla dikkat gerektirir çünkü değerlendirme öğretimin bir parçasıdır ve bilişsel ve sosyo-kültürel temelleri vardır. Bu olgulara dayanarak, bu vaka çalışması, üç İngilizce öğretmeninin, sınıflarındaki çocukların konuşma ve yazma becerilerini nasıl değerlendirdiklerine ilişkin görüş ve uygulamalarını araştırmaktadır. Veriler yarı yapılandırılmış görüşmeler yoluyla toplanmış ve içerik analizine tabi tutulmuştur. Bulgular öğretmenlerin görüş ve uygulamaları arasında uyuşmazlıklar olduğunu göstermektedir. Öğretmenlerin, öğretmen yetiştirme ve geliştirmedeki problemler, konuşma ve yazma becerilerin değerlendirilmesinden kaçınılması ve puanlayıcı eğitimi gerekliliği gibi çeşitli nedenlerden dolayı değerlendirici kimliğini benimsemedikleri bulunmuştur. Bulgular öğretmenlerin değerlendirme okuryazarlığının iyileştirilmesi gerektiğini göstermektedir. Bu nedenle, öğretmen yetiştirme programları, yabancı dil öğretmenlerinin değerlendirici kimlik geliştirmeleri üzerinde durmalıdır. Ayrıca yabancı dil öğretmenlerinin değerlendirme okuryazarlıklarını geliştirmek için hizmet içi eğitimler yapılmalıdır.

Anahtar kelimeler: Değerlendirme okuryazarlığı, yabancı dilde değerlendirme, genç öğrenenler, yazmayı değerlendirme, konuşmayı değerlendirme.

A teacher is sine qua non of classroom-based language testing and assessment (LTA). Teachers need to assess the success and performance of their students for making educational decisions through their teaching lives, and this assessment occurs in various situations, for example, when developing teaching materials, arranging courses, adapting the pace of instruction, managing the classroom, selecting homework, providing feedback as well as deciding on scores, placement, and monitoring (e.g., Allal, 2013; Glock, Krolak-Schwerdt, & Pit-ten Cate, 2015; Stiggins, 1991; Thiede et al., 2015). Assessment related activities are reported to result in teachers spending half of their time (Plake, 1993); therefore, teachers require being assessment literate so as to ensure time management.

The literature abounds with definitions related to language assessment literacy which was rooted in the term *assessment literacy* coined by Stiggins (1991). Stiggins (1995) advocates that teachers who are good at assessment recognise things to be assessed, the reason for performing it, the appropriate way to gauge the ability, knowledge of interest, how to create favourable instances of learner performance, things that are not suitable for evaluation, and the way to block that from occurring). Teachers' LTA literacy consists of various skills that help the individual to comprehend, assess as well as generate language tests and analyse test data (Pill & Harding, 2013). O'Loughlin (2013) views it as a series of competencies "related to testing production, test score interpretation and use, and test evaluation in conjunction with the development of a critical understanding about the roles and functions of assessment within society" (p.363). Fulcher (2012) asserts that the LTA literacy of teachers refers to the knowledge and abilities that a person should have for planning, developing, maintaining or assessing, large-scale standardised, and/or classroom-based tests.

Teacher assessment literacy plays a vital role in deciding on the quality of classroom assessments. Assessment is regarded as the crucial thing we can do to help our students learn as it both initiates and fosters learning (White, 2009). Teachers' competence in LTA has remained under-investigated (Davison & Leung, 2009; Fulcher, 2012; Tsagari & Csépes, 2011) because recent research has delved into the vital role of assessment in student learning – both the impact of important examinations and the need for qualified practice-based classroom assessment (Looney, Cumming, van Der Kleij, & Harris, 2017). A considerable number of studies display that EFL teachers generally lack sufficient assessment literacy and that their classroom assessment practices stay at the alarming rate (Ölmezer-Öztürk & Aydın, 2019; Tsagari & Vogt, 2017; Vogt & Tsagari, 2014). In the Turkish context, Hatipoğlu (2015) conducted a study with 124 student teachers to examine the assessment knowledge of pre-service teachers and their expectations from the testing course. The researcher found that all of the students expected to assess, choose and create tests, and prepare their learners for all types of tests. The pre-service teachers were also noted to have had limited assessment knowledge. Öz and Atay (2017) investigated Turkish EFL instructors' views on in-class language assessment and its link with their classroom practices. Findings revealed that although instructors claimed to be familiar with the basic terms related to classroom assessment, they had difficulty in reflecting their assessment knowledge into their classroom practice. Mede and Atay (2017) examined the assessment literacy of 350 Turkish EFL teachers working at the preparatory schools through an online questionnaire and found that participating teachers had limited assessment literacy and

need training in many domains of testing and assessment. Ölmezer-Öztürk and Aydın (2019) conducted a study in higher education with 542 EFL teachers working at schools of foreign languages to provide an overall picture regarding their general and skill-based Language Assessment Knowledge (LAK) level. They found that teachers had insufficient language assessment knowledge. Finally, Şişman and Büyükkarcı (2019) reviewed studies (from 1987 to 2019) examining language teachers' language assessment literacy and warned that the assessment literacy level of language teachers is low and need to be developed.

As shown in the research mentioned above, studies related to language assessment literacy of EFL teachers, particularly in the Turkish context, were conducted with adults (in higher education on this issue). Thus, there is an urgent need for identifying the LTA literacy of Turkish ELF teachers working at primary and secondary education levels. Primary schools are one of the workplaces in which teachers are asked to teach English, and gauge their students in every skill and integrated skills. In this level, teachers are expected to have the required knowledge of assessment as well as the development of learners. As a starting point, it is a must to contribute to the literature in terms of determining the LTA literacy of Turkish ELF teachers of young learners. In light of these observations, to address the gaps in previous research, the current study aims to gather in-depth information about the LTA literacy of Turkish ELF teachers of young learners.

Teacher Assessment Literacy in Practice (TALiP)

Based on relevant literature in teacher education, Xu and Brown (2016) have developed TALiP as a conceptual framework, which gains insights from Willis, Adie, and Klenowskis' (2013) model and DeLuca's (2012) framework (see Figure 1). This conceptual framework includes six components, starting from the knowledge base which is at the bottom of the pyramid, and ending with teacher identity (re)construction as assessors which stays at the top of the pyramid.

The *knowledge base remains* at the bottom of this pyramid. Xu and Brown (2016) advocate the knowledge base as a necessary, but insufficient condition in terms of TALiP, although no standards or criteria would exist without the knowledge base; therefore, causing poor or incorrect outcomes for teachers and students (Fulcher, 2012). This component consists of seven kinds of knowledge, and it can be conceded as a threshold. Teachers need to cross it in order to gain a chance to engage in assessment at a deeper level. The second component is about teacher understanding of assessment which is considered to be crucial for filtering and interpreting knowledge (Barnes, Fives, & Dacey, 2015; Fives & Buehl, 2012). Teacher beliefs, according to Xu and Brown (2016), serve as an explanatory and guiding framework through which theoretical knowledge is acquired and applied. According to them, teachers' understanding of assessment includes cognitive and affective aspects. While the former shows teachers' beliefs about what is right and false about assessment, the latter refers to emotional tendencies that teachers display concerning different facets and uses of assessment.
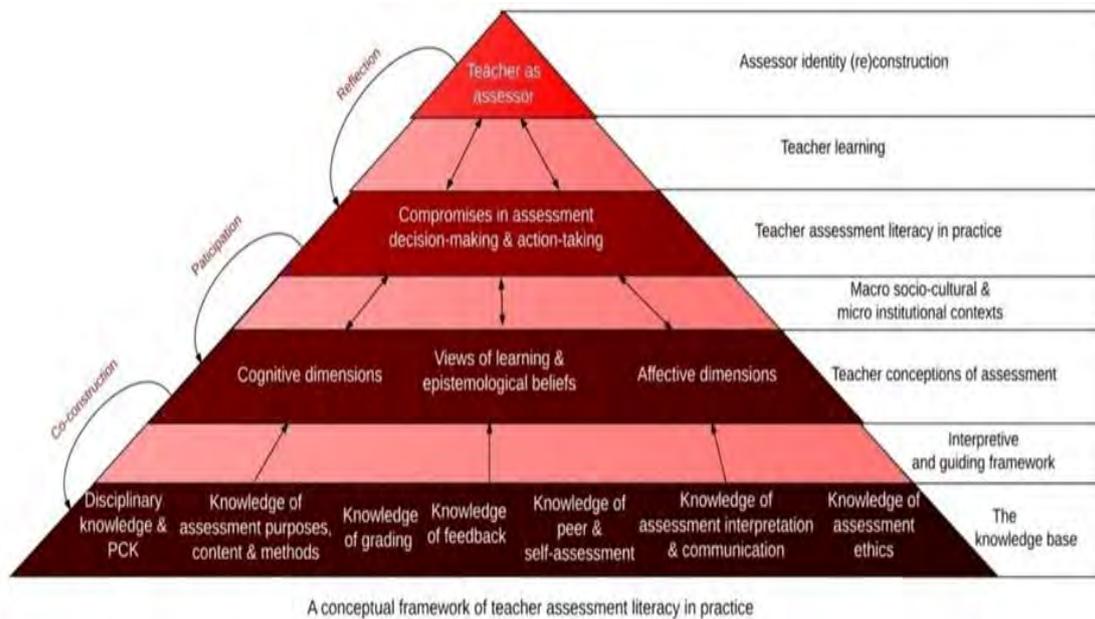
A conceptual framework of teacher assessment literacy in practice

Figure 1. Xu and Brown's (2016) Conceptual Framework of TALiP

The third component indicates that working in a common place and broader communal, political, and cultural contexts limit in-service teachers act whatever they please in real performance. These variables affect teachers' assessment practices via policies and rules to generate a habit of certainty and consent that is not easily disputed by practitioners (Scarino, 2013). These boundaries can be in various forms, being small as pre-specified criteria or being large as nationwide assessment rules (Xu & Brown, 2016). The aforementioned contextual variables establish boundaries, which determine what teachers should do and avoid in their assessment practices (Gu, 2014). The fourth component refers to teachers' efforts to adjust the requirements of extrinsic determinants and their own beliefs (McMillan, 2003). Therefore, Xu and Brown (2016) argue that assessment literacy is better recognised as TALiP, which includes different compromises that teachers make to appease tensions. The fifth component of TALiP aims to help teachers give up on imitating traditional exercises that do not go with effective practices and foster teacher learning. From this point of view, teacher learning can be accepted as the driving force to influence changes in assessment and promote TALiP. The final component serves as the ultimate goal of TALiP. The teacher is considered as an assessor of learning as well as an instructor. If teachers become aware of their own identity as assessors, they likely become assessors of their own practices and they can combine different perspectives into their own understandings (Xu & Brown, 2016). This identity (re)construction as assessors helps them make more substantiated agreements in their assessment applications that may result in enhanced TALiP.

Assessing Young Language Learners

Assessment has many goals, and one of them is to help teachers discover how much their learners have learned throughout the process. Regarding the assessment of young language learners, recent research has focused on the range as well as the value

of teacher assessment and the evaluation process within formative and summative assessment (Rea-Dickins & Gardner, 2000). The results obtained by Rea-Dickins and Gardner (2000) indicate that teachers may need to utilise the formal tests to investigate the language targets that young learners may be considered to succeed since using informal classroom tests is not always appropriate. Although Chou (2014) used games, songs and stories to ease children's learning and studying of English words in her research, she underlined the necessity of adopting formal tests to investigate what children had acquired by the end of the lesson. Testing young learners does not resemble assessing adult language learners. Considering testing young learners, Hughes (2003) and McKay (2006) made several recommendations, such as using short and diverse tasks, and pictures. Teachers are expected to incorporate various assessment practices because using only one assessment method makes it impossible to meet the needs of students (Stiggins, 2002). In addition, Earl (2003) insists that teachers should avoid using only one assessment practice. Instead, they need to apply the amalgamation of assessment for (formative assessment), of (summative assessment), and as learning. The common point in all these statements is that the teacher needs to be equipped with various purposes of assessment and be able to utilise them accordingly (Green & Mantz, 2002). The present study aimed to identify the LTA literacy of Turkish ELF teachers in terms of their views and practices; to determine their assessment types; to examine their teacher identity. Accordingly, the research questions of the present study are as follows:

1. What are the views and practices of Turkish EFL teachers on language assessment?

2. What is the perceived identity of Turkish EFL teachers? (Teacher as an instructor? Teacher as an assessor? Or Both?)

Method

The present study follows a qualitative multiple case studies approach because it allows an in-depth analysis of cases and collects detailed information using various data collection procedures over a while (Creswell, 2013). The multiple cases also help to understand the situation by studying similarities and differences among the cases. In a case study, "a how or why question is being asked about a contemporary set of events over which the investigator has little or no control" (Yin, 1994, p. 9). The researchers seek answers to *how* and *why* questions regarding the assessment beliefs and practices of foreign language young learner teachers.

Participants

Three EFL teachers participated in the study. The participants were recruited purposively as a part purposive sampling, which "is a practical and efficient tool when used properly, and can be just as effective as, and even more efficient than, random sampling" (Tongco, 2007, p. 155), and they were the teachers of 5th-grade prep-class students in the academic year of 2017-2018 (please see Table 1 for their demographic characteristics). Before participating in the study, all participants provided written informed consent for ethical concerns.

Table 1

Demographics of Participants

Teacher	Gender	Age	Years of Experience	Graduation Degree*
Beril	Female	24	2	BA
Ali	Male	40	15	BA
Metin	Male	28	5	MA

*BA: Bachelor's degree, MA: Master's degree.

Table 1 presents that three EFL teachers (one female and two males) participated in the study. Their ages ranged from 24 to 40. They had been teaching English for 2, 5, and 15 years. Two of the teachers had a BA, while one had an MA. Although it was not indicated in the table, the data obtained from the interviews showed that teachers were working in the eastern part of Turkey and graduated from English language teaching (ELT) programs. There are two primary foreign language teacher education programs in Turkey. The first one is ELT programs of educational faculties. These programs are at the undergraduate level and earning a bachelor's degree requires four years of education. A second option is ELT certificate programs for graduate or ongoing students of linguistics, British language and literature, translation and interpreting or American culture and literature programs are one other way of becoming English teachers.

Context

In Turkey, primary school students begin learning English in the second grade. Recently, fifth grade has been determined as an intensive-English class within a project devised by the MoNE. The academic year of 2017-2018 was determined as the piloting year for the project, which was implemented in 620 secondary schools across Turkey. The project focuses on teaching English to young learners throughout a well-structured and balanced intensive English program (MoNE, 2017a). In other words, students at the age of 11 and 12 were provided with an intensified English class who had limited opportunities to use the target language in real life outside the classroom. This issue is common in countries like Turkey in which English is a foreign language. The MoNE (2017a) asserts that the program follows the tenets of communicative approach and it has been designed in line with the principles of CEFR by attaching significance to learner autonomy, critical thinking, and problem-solving. The participants of this study were graduates of ELT programs. The program includes more cultural elements to enhance students' cultural awareness and intercultural communication skills (Dincer & Koç, 2020).

Data Collection and Analysis

The data were gathered via semi-structured interviews. A semi-structured interview is a more flexible and common version of interviews (Holloway & Wheeler, 2010). Rubin and Rubin (2005) advocate that a semi-structured interview "allows depth to be achieved by providing the opportunity on the part of the interviewer to probe and expand the interviewee's responses" (p. 88). The researchers of the current study formed

a semi-structured interview forms consisting of two parts. The first part consisted of questions about participants' demographic information (gender, the region they work, professional experience, university … etc.). In the second part, questions and sub-questions aimed to identify their LTA literacy. The interview process followed two phases. In the first phase, the researchers made contact with participants to give brief information about the aim and content of the study. The goal of this step is to make interviewees familiar with the research environment for ensuring credibility (Richards, 2003). To increase both quantity and quality of the data, teachers were interviewed in their native language (Mackay & Gass, 2005). These interviews were audiotaped and transcribed.

Data analysis was carried out cyclically, with analysis being done on one batch of data as a way to inform the next steps of the research. To discover, develop, and form well-grounded interpretations, systematic data coding process is of importance (Mackey & Gass, 2012). While analysing the data, the researchers followed Dörnyei's (2007) stages of content analysis. First, all data were transcribed, because transcription is the first step of analysing and discovering themes (Rubin & Rubin, 2012). To ease the analysis process, the researcher adopted 'edited transcription' during transcribing the data, omitting hmm, uh, huh, etc. (Hansen, 2003, p. 136). Second, the researcher read the data for getting a general idea and for checking pre-coding. Then, the data was coded and categorised under themes by making interpretations and collaborative discussions were conducted to enrich the interpretation. The remaining data analysis procedure involved negotiating agreement for discrepancies in coding. Additionally, a third coder served as an external auditor (Creswell, 2012). Miles and Huberman's (1994) formula was applied for interrater reliability which was found as 90%, which is a satisfactory level (Miles & Huberman, 1994). The analysis utilised peer debriefing and member check to increase reliability. Pseudonyms were used in reporting the findings.

Results

Findings were provided with respect to research questions. As for the first research question, which aimed at displaying the views and practices of Turkish EFL teachers on language assessment, the findings of the semi-structured interviews were given below. Table 2 displays not only the views and practices of English language teachers in relation to language assessment but also tracks the harmony of the views and practices.

Table 2

Turkish EFL Teachers' Views and Practices on Language Assessment

Teachers	Views	Practices
Beril	Multiple assessment tools and techniques are required	Grammar, vocabulary, and translation
Ali	Productive skills must be assessed	Grammar-based exams
Metin	Presentations, tasks, demonstrations, role-plays	Grammar-based exams, and speaking

As Table 2 shows, there are mismatches between teachers' views and practices. More specifically, they had difficulty in reflecting their assessment beliefs and knowledge into their classroom practice. For example, in terms of views, participants believed that;

> **Beril**: It is not ideal to assess learners' language proficiency only via traditional paper-pen based assessment tools. As far as I know, language proficiency cannot be gauged appropriately through multiple-choice questions. Instead, communicative activities and speaking sessions are of importance in addition to traditional assessment tools.

> **Ali**: Great importance needs to be paid to speaking and writing skills. Students write every word in English as they were pronounced due to the structure of their native (Turkish) language. Thus, it is important to foster their writing and speaking as well as pronunciation skills.

> **Metin**: Productive skills also need to be assessed. For example, students can be given some argumentative topics and they may be asked to present or demonstrate the related role-plays.

Although teachers underlined the importance of using various assessment tools and techniques and assessing productive skills through presentations, demonstrations, and role-plays, they were found to apply mostly grammar-based exams. They showed students' low level of language proficiency as one of the reasons for this situation. The other issue was claimed to be the fact that assessing productive skills is not a must. Here are some representative statements:

> **Beril**: As I mentioned before, I disagree with the traditional assessment tools. However, students are not good at speaking skills. They cannot even spell their names in English. Therefore, it is inevitable for me to conduct exams based on grammar, vocabulary, and translation although I tried to include listening and writing activities.

> **Ali**: Instead of taking into account the productive skills, the scores of paper-pen based exams are used for student evaluation. In other words, since it is not compulsory, I do not assess students through speaking activities.

> **Metin**: We are asked to use paper-pen based exams through for giving scores. Nevertheless, I sometimes gauge students' speaking ability and take into account their speaking scores.

The second research question examined whether Turkish EFL teachers were aware of their own identity as assessors. The findings showed that teachers did not adopt the assessor identity due to various barriers; namely, *problems in teacher training and development, avoiding assessing productive skills*, and *the necessity for rater training* (see Table 3). For example, the participants tended to use familiar traditional textbook-based assessment tools. They did not prefer using core speaking and writing activities. Moreover, they had low confidence in assessing speaking and writing skills. Last, they had difficulties in selecting the appropriate rubric or preparing their own rubrics.

Table 3

The Barriers to Teacher Assessment Literacy in Practice

Themes	Sub-themes
Problems in teacher training and development	• Lack of courses on assessing young learners • Need for in-service teacher education
Avoiding assessing productive skills	• Mere focus on assessing grammar and vocabulary • Avoiding assessment of communicative skills
The necessity for rater training	• Preparing rubrics • Using rubrics • Providing feedback on assessment • Skill-based assessment.

In the first theme concerning the problems in teacher training and development, three issues are addressed*: lack of courses on assessing young learners* and *the need for in-service teacher education*. All of the participants emphasised the related changes in both pre-service and in-service training considering the effective assessment of young learners.

> **Beril**: In my pre-service education, I do not remember whether we focused on assessing young learners or not. It would be better to learn how to assess young learners.

> **Ali**: I do not know how to assess the productive skills of young learners appropriately. Although there are available rubrics for serving this purpose, our scores for the same performance differ from each other, because we need support not only about the assessment of young learners but also about their development (cognitive, affective, physical, etc.).

Unlike Beril and Ali, Metin seemed to utilise speaking activities and assessment of speaking as he stated that "*Although we are asked to assess students through paper-pen based exams for giving scores, I also gauge students' speaking ability and take into account the speaking scores*." However, the data showed that he used English only on one day of the week. He confessed that "*On Fridays, we force ourselves to use only English.*"

Avoiding assessing productive skills is the second theme that confirms the result of the first research question, indicating that there are mismatches between teachers' views and practices. This means that teachers pay attention to the assessment of grammar and vocabulary but avoid assessing communicative skills. Here are some examples:

> **Beril**: The focus is on assessing grammar, vocabulary and translation. I did not assess speaking and writing skills.

> **Ali**: Although we are trying to include assessing speaking next semester, now our exams are grammar oriented.

The last theme, the necessity for rater training, showed the necessity for training on preparing rubrics, using rubrics, providing feedback on assessment, and skill-based assessment. Here are the views of the teachers:

> **Beril**: I would prefer we had an opportunity for training on the assessment of young learners as it is supposed to be more communicative. You know, it is got to be paired with games.

> **Ali**: I would like to attend the INSET programs by efficient professors in order to compensate for my inadequacy in scoring the performance-based activities.

Metin: I believe I would benefit from a rater training in various ways, such as preparing my own speaking rubric.

According to this finding, teachers are in need of receiving rater training programs because they consider themselves inadequate in terms of preparing and using rubrics. In addition, teachers think that such kind of rater training programs may help them score the performance-based activities appropriately.

Discussion

Regarding the first research question, various mismatches occurred between teachers' views and practices. In other words, teachers were observed to have difficulty in reflecting their assessment beliefs and knowledge into their classroom practice. This confirms the findings noted by Öz and Atay (2017), who examined the relationship between perceptions and practices of Turkish EFL instructors towards in-class language assessment. They found an imbalance between teachers' assessment literacy and their classroom reflection, although the majority asserted to be acquainted with basic classroom assessment. In other words, their research did not reveal enough connections between the experience and assessment perception. In addition, Muñoz, Palacio, and Escobar (2012) investigated teacher's beliefs of assessment in general and whether they could put those beliefs into practice or not. Their study demonstrated a mismatch between teachers' beliefs and their practices. Xu and Brown (2016) assert that teacher beliefs serve as an interpretive and guiding framework through which theoretical knowledge is acquired and applied. Although teachers considered multiple assessment tools and techniques as vital for effective assessment of young learners, they mainly assessed their students via grammar-based exams. The study findings provide insight into the culture of language teaching in Turkey. Instead of communicative activities, Turkish teachers were reported to use the grammar-translation method (Büyükyavuz & İnal, 2008; Oktay, 2015). Moreover, Coskun (2016, p. 1), investigating "I can understand English, but I can't speak" syndrome in Turkey, found that some of the reasons for this syndrome are the grammar-based teaching in English lessons, inadequate speaking practices, speaking anxiety, and use of L1 by the teacher. The current study confirms this fact. Even though the course books and materials were communicatively oriented, the teachers of the young learners in this study had to focus on the forms of the language as assessing forms in the language takes less time. These mismatches may be due to the teachers' efforts to balance the requirements of outside factors and their own beliefs (McMillan, 2003).

The second research question was about the teacher assessment identity. That is, it questioned whether teachers considered themselves as assessors or not. If teachers do not consider themselves as assessors, what are the barriers to the teacher assessor identity? MoNE has recently redefined general teacher competencies under three main components: 1) teacher knowledge, 2) teacher skills, and 3) attitude and values (MoNE, 2017b). Teacher skills pertain to in- and out-of-class practices to design and manage teaching and assessment. Therefore, assessment is one of the skills that are necessary for teacher competency. A teacher is considered as an assessor of learning as well as an instructor. If teachers become aware of their own identity as assessors, they become the creators of their own evaluation practices, started doing self-reflection, and able to combine others' perspectives with their own values (Xu & Brown, 2016). However, it was seen that teachers did not feel as assessors because of various obstacles. For

example, although the English Teacher Education program (ELTE) focuses on training young learner EFL teachers, *Ali* does not remember a specific focus on assessing young learners. He believes he did not put into his pockets anything regarding the assessment of young learners. This finding can be claimed to be an urgent call for some continuing professional development sessions to the English language teaches to increase their awareness as well as to enhance their practices about their own assessment practices. As underlined by Stiggins (1988), much time and effort are needed for classroom assessment, because assessment-related activities in which teachers directly involved consume 40% of their time. It is also believed that there should be at least one assessment course in any effective pre or in-service program involving practitioners with evaluation expertise as well as practical experience (Kahl, Hofman, & Bryant, 2012).

Another issue was the avoidance of assessing communicative skills. Although participating teachers believed that young learners easily got bored during grammar-based instruction, they generally avoided assessing speaking. Young learners complied with their inappropriate pre- and in-service education, so the participants seem to follow the same routines with other teachers, instead of being the authors of their own assessment practices. This is also related to teachers' culture of learning. More specifically, the teachers' experiences in their own language learning seem to affect their teaching. The findings showed that the teachers focused on assessing the language forms instead of the language communicative functions. In other words, the teachers approach the language as an object but not teach the language itself. In other words, teachers try to teach about the language (e.g., grammar rules) instead of teaching the language itself (e.g. teaching the functions). The functional guidance offered in the Common European Reference for Languages (CEFR). In the CEFR (Council of Europe, 2001), the descriptors refer to communication practices demonstrated by language learners who are prone to perform the language for real-life purposes (Benigno & de Jong, 2016). Moreover, when the beliefs and practices of the participants are considered, it is reasonable to conclude that the teachers should be trained and motivated to utilise authentic assessment. The findings show that the teachers are willing to assess with alternative tools, but they are not motivated and not trained to do so. Pre-service and in-service teacher education should help teachers use alternative tools such as drama since it can be used to enrich and enliven the assessment (Rothwell, 2012).

The last theme (the necessity for rater training) showed the necessity for training on preparing rubrics, using rubrics, providing feedback on assessment, and skill-based assessment, which can be useful on reliable and dependable results as analytic and holistic scoring (Han, 2013). Considering Xu and Brown's (2016) TALiP, it is reasonable to conclude that the teachers lack sufficient information almost in all levels of the framework. Thus, as underlined by the participants, teachers need to be provided opportunities for training on assessment through the INSET programs by efficient professors in order to foster the quality of teacher assessment. Turkish EFL teachers in a study conducted by Mede and Atay (2017) asked for training in productive skills particularly speaking.

Conclusion

This study investigated the assessment literacy of three Turkish ELF teachers regarding their views and practices. The findings revealed that EFL teachers' assessment of young learners is not yet satisfying. There were some discrepancies between teachers' views and practices of the participating teachers. In other words, the participants could not reflect their assessment beliefs and knowledge into their classroom practice. The findings also showed that the teachers did not adopt the assessor identity due to various barriers; namely, problems in teacher training and development, avoiding assessing productive skills, and the necessity for rater training. When the final goal of TALiP (Xu & Brown, 2016) is considered, the findings of this research show that applying this model can be a good way to improve the assessment skills of language teachers. Considering the results of this current study, it is fair to advocate that the teachers of young learners should be provided with efficient INSET.

Implications and limitations

The study has some implications for stakeholders. First, the findings suggest that the assessment literacy of the teachers needs to be improved. Therefore, pre-service teacher education programs should emphasise the identity development of language teachers as assessors. Second, for in-service teachers, effective rater training INSET program should be initiated. Pre- and in-service education are essential because educating pre- and in-service teachers establish one of the most crucial dimensions in the quality assurance of language testing and assessment (Vogt & Tsagari, 2014). Third, school principals should provide language teachers with support for implementing alternative assessments by allowing them to be flexible in the assessment. Last, language teachers should notice the significance of the assessment. The teachers can use functional guidance suggested in the CEFR. More specifically, the age-appropriate set of functional descriptors of Benigno and de Jong (2016) can be used to assess young learners.

The study has some limitations. One limitation is that it was conducted with three EFL teachers and the teachers' self-reported data may be subject to bias. Another limitation is the lack of a diverse teacher population. Considering the findings and the limitations of this present study, it can be concluded that there is a need for more studies on assessment literacy of primary and secondary school teachers in Turkey. A future study can observe language teachers' classroom assessment practices. More in-depth large-scale studies can be carried out by using several data collection procedures such as observing classrooms or gathering data from various sources such as teacher and learner journals, narratives and the results of needs analyses. In addition, a similar study can be conducted with more participants from various backgrounds or schools so that it can indicate the implications of the larger groups. Last, further research studies can consider including schools from different regions in Turkey.

Statement of Responsibility

Sabahattin Yeşilçınar; conceptualisation, methodology, validation, formal analysis, writing–original draft, writing-reviewing & editing, visualisation, supervision, and project administration. Galip Kartal; conceptualisation, methodology, validation, investigation, data curation, writing–original draft, writing-reviewing & editing.

References

Allal, L. (2013). Teachers' professional judgement in assessment: A cognitive act and a socially situated practice. *Assessment in Education: Principles, Policy and Practice*, *20*(1), 20-34. https://doi.org/10.1080/0969594X.2012.736364

Barnes, N., Fives, H., & Dacey, C. M. (2015). Teachers' beliefs about assessment. In H. Fives & M. G. Gill (Eds.), *International Handbook of Research on Teacher Beliefs* (pp. 284-300). New York: Routledge.

Benigno, V., & de Jong, J. (2016). The global scale of English learning objectives for young learners: A CEFR-based inventory of descriptors. In M. Nikolov (Ed.), *Assessing young learners of English: Global and local perspectives*. New York: Springer.

Büyükyavuz O., & İnal S. (2008). A descriptive study on Turkish teachers of English regarding their professional needs, efforts for development and available resources. *The Asian EFL Journal*, *10*(3), 215-234.

Chou, M. H. (2014). Assessing English vocabulary and enhancing young English as a foreign language (EFL) learners' motivation through games, songs, and stories. *Education 3-13*, *42*(3), 284-297.

Coskun, A. (2016). Causes of the 'I can understand English but I can't speak' syndrome in Turkey. *i-manager's Journal on English Language Teaching*, *6*(3), 1-12.

Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge, UK: Cambridge University.

Creswell, J. W. (2012). Educational research: Planning, conducting, and evaluating quantitative and qualitative research. Boston, MA: Pearson Publication.

Creswell, J. W. (2013). Research design: Qualitative, quantitative, and mixed methods approaches, Sage publications.

Davison, C., & Leung, C. (Eds.) (2009). Teacher-based assessment. *TESOL Quarterly*, Special Issue *43*(3), 389–562.

DeLuca, C. (2012). Preparing teachers for the age of accountability: Toward a framework for assessment education. *Action in Teacher Education*, *34*(5/6), 576-591.

Dincer, A., & Koç, H. K. (2020). The implementation of an intensive English language program in the fifth grade in Turkey: A qualitative evaluation. *Journal of Theoretical Educational Science, 13*(1), 25-43.

Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative and mixed methodologies*. Oxford: Oxford University.

Earl, L. (2003). *Assessment as learning: Using classroom assessment to maximise student learning*, Corwin Press, Thousand Oaks, CA.

Fives, H., & Buehl, M. M. (2012). Spring cleaning for the "messy" construct of ceachers' beliefs: What are they? Which have been examined? What can they tell us? In K. R. Harris, S. Graham and T. Urdan (Eds.), *APA educational psychology handbook: Individual differences and cultural and contextual factors* (Vol. 2, pp. 471-499). Washington, DC: American Psychological Association.

Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, *9*(2), 113-132.

Glock, S., Krolak-Schwerdt, S., & Pit-ten Cate, I. M. (2015). Are school placement recommendations accurate? The effect of students' ethnicity on teachers' judgments and recognition memory. *European Journal of Psychology of Education, 30*, 169-188.

Green, S. K., & Mantz, M. (2002). Classroom assessment practices: Examining impact on student learning. Paper presented at the *Annual Meeting of the American Educational Research* Association, New Orleans, LA.

Gu, P. Y. (2014). The unbearable lightness of the curriculum: What drives the assessment practices of a teacher of English as a Foreign Language in a Chinese secondary school? *Assessment in Education: Principles, Policy and Practice*, *21*(3), 286-305.

Han, T. (2013). *The impact of rating methods and rater training on the variability and reliability of EFL students' classroom-based writing assessments in Turkish universities: An investigation of problems and solutions* [Doctoral dissertation]. Atatürk university, Erzurum, Turkey.

Hansen, K. (2003). *Writing in the social sciences: A rhetoric with readings*. Boston: Pearson Custom.

Hatipoğlu, Ç. (2015). English language testing and evaluation (ELTE) training in Turkey: Expectations and needs of pre-service English language teachers. *ELT Research Journal*, *4*(2), 111-128.

Holloway, I., & Wheeler, S. (2010). *Qualitative research in nursing and healthcare*. Oxford, UK: Blackwell.

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

Kahl, S. R., Hofman, P., & Bryant, S. (2012). Assessment literacy standards and performance measures for teacher candidates and practicing teachers. Measured Progress. *Assessment*. Retrieved from https://www.memphis.edu/caep/caep/assessment-literacy-in-teacher-preparati.pdf.

Looney, A., Cumming, J., van Der Kleij, F., & Harris, K. (2018). Reconceptualising the role of teachers as assessors: teacher assessment identity. *Assessment in Education: Principles, Policy & Practice*, *25*(5), 442-467.

Mackay, A., & Gass, S. (2005). *Second language research: Methodology and design*. mahwah, New Jersey: Lawrence Erlbaum Associates.

Mackey, A., & Gass, S. (2012). *Research methods in second language acquisition: A practical guide*. Oxford: Wiley-Blackwell

McKay, P. (2006). *Assessing young language learners*. Cambridge: Cambridge University.

McMillan, J. H. (2003). Understanding and improving teachers' classroom assessment decision making: Implications for theory and practice. *Educational Measurement: Issues and Practice*, *22*(4), 34–43.

MoNE (2017a). Yabancı dil ağırlıklı 5. Sınıf İngilizce dersi öğretim programı. Retrieved from

https://dogm.meb.gov.tr/meb_iys_dosyalar/2017_09/26153336_5._SINIF_YNG._P ROGRAMI.pdf

MoNE (2017b). Öğretmenlik mesleği genel yeterlikleri. Ankara: Öğretmen Yetiştirme ve Geliştirme Genel Müdürlüğü. Retrieved from http://oygm.meb.gov.tr/meb_iys_dosyalar/2017_12/11115355_YYRETMENLYK_ MESLEYY_GENEL_YETERLYKLERY.pdf

Mede, E., & Atay, D. (2017). English Language Teachers' assessment literacy: The Turkish context. *Dil Dergisi, 168*(1), 1-5.

Miles, M. B., & Huberman, A. M. (1994). Qualitative data analysis: An expanded sourcebook. London, UK: Sage

Muñoz, A. P., Palacio, M., & Escobar, L. (2012). Teachers' beliefs about assessment in an EFL context in Colombia. *Profile Issues in Teachers Professional Development, 14*(1), 143-158.

Oktay, A. (2015). Foreign language teaching: A problem in Turkish education. *Procedia - Social and Behavioral Sciences*, *174*(1), 584-593.

O'Loughlin, K. (2013). Developing the assessment literacy of university proficiency test users. *Language Testing*, *30*(3), 363-380.

Ölmezer-Öztürk, E., & Aydın, B. (2019). Investigating language assessment knowledge of EFL teachers. *Hacettepe University Journal of Education*, *34*(3), 602-620. https//doi.org/10.16986/HUJE.2018043465

Öz, S., & Atay, D. (2017). Turkish EFL teachers' in-class language assessment literacy: Perceptions and practices. *ELT Research Journal*, *6*(1), 25-44.

Pill, J., & Harding, L. (2013). Defining the language assessment literacy gap: Evidence from a parliamentary enquiry. *Language Testing, 30*(3), 381-402.

Plake, B. S. (1993). Teacher assessment literacy: Teachers' competencies in the educational assessment of students. *Mid-Western Educational Research, 6*(1), 21-27.

Rea-Dickins, P., & Gardner, S. (2000). Snares and silver bullets: Disentangling the construct of formative assessment. *Language Testing*, *17*(2), 215-243.

Richards, K. (2003). *Qualitative inquiry in TESOL*. Springer.

Rothwell, J. (2012). Drama and languages education: authentic assessment through process drama. In Winston, J (Ed.), *Second language learning through drama: practical techniques and applications* (pp. 54-68). Routledge, United States of America.

Rubin, H. J., & Rubin, I. S. (2005). *Qualitative interviewing: The art of hearing data* (2nd ed.). Thousand Oaks, CA: Sage.

Rubin, H. J., & Rubin, I. S. (2012). *Qualitative interviewing: The art of hearing data.* Los Angeles: SAGE.

Scarino, A. (2013). Language assessment literacy as self-awareness: Understanding the role of interpretation in assessment and in teacher learning. *Language Testing, 30*(3), 309-327.

Stiggins, R. J. (1988). Revitalising classroom assessment: The highest instructional priority. *The Phi Delta Kappan*, *69*(5), 363-368.

Stiggins, R. J. (1991). Assessment literacy. *Phi Delta Kappan*, *72*(7), 534-539.

Stiggins, R. J. (1995). Assessment literacy for the 21st century. *Phi Delta Kappan*, *77*(3), 238-245.

Stiggins, R. J. (2002) Assessment crisis! The absence of assessment FOR learning. *Phi Delta Kappan, 83*(10), 758-765.

Şişman, E. P., & Büyükkarcı, K. (2019). A review of foreign language teachers' assessment literacy. *Sakarya University Journal of Education*, *9*(3), 628-650.

Thiede, K. W., Brendefur, J. L., Osguthorpe, R. D., Carney, M. B., Bremner, A., Strother, S.,…Jesse, D. (2015). Can teachers accurately predict student performance? *Teaching and Teacher Education*, *49*, 36-44.

Tongco, M. D. C. (2007). Purposive sampling as a tool for informant selection. *Ethnobotany Research and Applications*, *5*, 147-158.

Tsagari, D., & Csépes, I. (2011) (Eds.). *Classroom-based language assessment*. Frankfurt am Main: Peter Lang.

Tsagari, D., & Vogt, K. (2017). Assessment literacy of foreign language teachers around Europe: Research, challenges and future prospects. Papers in *Language Testing and Assessment*, *6*(1), 41-63.

Vogt, K., & Tsagari, D. (2014). Assessment literacy of foreign language teachers: Findings of a European study. *Language Assessment Quarterly, 11*(4), 374-402.

White, E. (2009). Are you assessment literate? Some fundamental questions regarding effective classroom-based assessment. *OnCUE Journal, 3*(1), 3-25.

Willis, J., Adie, L., & Klenowski, V. (2013). Conceptualising teachers' assessment literacies in an era of curriculum and assessment reform. *The Australian Educational Researcher*, *40*(2), 241-256.

Xu, Y., & Brown, G. T. L. (2016). Teacher assessment literacy in practice: A Reconceptualisation. *Teaching and Teacher Education*, *58*, 149-162.

Yin, R. K. (1994). *Case study research: Design and methods* (2nd ed.). Thousand Oaks, CA: Sage.