Iranian EFL Teachers' Assessment Literacy Knowledge: The Impact of the Educational System on Teachers' Classroom Assessment Practices

Karim Rezagah[1]*

Abstract
English language teachers' assessment ability to assess all areas of pupils' learning is important for comprehending how learners advance across the curriculum and guiding them in their development (Livingston & Hutchinson, 2016). The educational system of Iran decided to shift from traditional teaching methodologies toward communicative approaches. This conversion would not turn into reality unless teachers could apply it in practice. However, teachers have certain needs such as expanding and broadening contextual-related knowledge, cooperating with their colleagues, and developing their assessment literacy (Coombe, Vafadar, & Mohebbi, 2020). To identify the teachers' understandings of, practices in, and challenges of assessment, 15 English language headteachers (English language teachers who are responsible for moderating the group of the teachers' activities) participated in the interviews of this study, followed by questionnaires for exploring teachers' needs. The interviews were coded and content analyzed independently by the researcher and an expert in assessment. The main themes and needs were derived from the interview analyses and are presented in eight pivots. The findings of the questionnaires manifested the priorities that teachers felt concerning assessment literacy and classroom-based assessment needs. The paper will discuss the findings concerning assisting teachers' professional development in assessment literacy. Implications are also provided.

Keywords: assessment; classroom assessment; educational system; teacher assessment literacy

1. Introduction

Developing teachers' assessment ability to assess learners for fostering their learning, including educational tradition and assessment culture, needs better comprehension of the complexity of and dynamic interactions among curriculum, assessment, and pedagogy, student and teacher learning needs, and the contexts of learning and teaching (Xu & Brown, 2016). To develop teachers' assessment literacy, that is the knowledge assessors need to possess for accomplishing the actions that are relevant to assessment (Inbar-Lourie, 2017), teachers' understanding of assessment and its interrelationship with other key factors including the wider policy context and the social, cultural, and professional contexts is required (Livingston & Hutchinson, 2016, p. 2).

---
[1] Center for Applied Language Studies, University of Jyvaskyla, Jyvaskyla, Finland.
Email: krezagah@student.jyu.fi

Several researchers have studied changes in educational systems, impacting teachers' understanding of classroom assessment and their assessment practices Luxia (2007) in China, Livingston & Hutchinson (2016) in Scotland, and Green (2016) in Japan are some of these studies. Luxia (2007), in her study which focused on the writing skill, asserts that due to the pressure of preparing students for high-stakes tests, the communicative context of teaching writing was neglected and it was not taught for communication, but test preparation and obtaining higher scores. Therefore, writing communicatively for real-life situations is ignored because it is not included in the exam (Luxia, 2007).

In a similar effort, the study that Green (2016) conducted in Japan, considered the arguments for the testing of spoken language skills and the contribution that the use of such tests might make to language education. He argues that while the educational system in Japan contemplates mandating spontaneous development of the four skills, the university entrance examination accentuates the writing skill. The dominance of these high-stakes tests, affects teachers' classroom activities, their approaches, and the ultimate objectives of the course.

The situation of teaching English in Iran has a lot in common with both of these contexts. The educational system has prescribed communicative language teaching for schools. However, due to the dominance of high-stakes tests, we are far from reaching this objective (see Firoozi, Razavipour & Ahmadi, 2019).

1.1. Assessment for/of Learning

Assessment *of* learning (*AoL*) is typically administered at the end of a unit or grading period and evaluates a student's understanding by comparing his or her achievement against a class-, district, or nationwide benchmark or standard while assessment *for* learning (*AfL*) assesses a student's comprehension and understanding of a skill or lesson during the learning and teaching process (Bennett, 2017, Van der Kleij & Cumming, 2017). AfL and AoL are used interchangeably with formative assessment and summative assessment respectively in this study.

AoL and AfL are two of the main assessment purposes which are widely used in the classroom environment and debated largely in the literature. In recent years, assessment standards have moved from AoL toward AfL (DeLuca, LaPointe-McEwan & Luhanga, 2016). According to the Iranian context and the importance of the grades in reporting the students' achievement, exercising AoL and AfL harmoniously could yield more effective results, as Hildén and Fröjdendahl (2018), put it nicely "today, harmony between formative and summative assessment strategies is needed to ensure sufficient quality of various kinds of assessment across educational systems" (p. 2).

To further differentiate the two purposes of assessment, it can be stated that AoL assessments summarize the students' achievements from the course. The results could also be used for evaluating the effectiveness of the instruction. Summative tests are marked according to a scale or a set of criteria (Mousavi, 2009). Different from AoL, AfL is student-centered and encourages learners to shoulder the responsibility for their learning process. The teacher's responsibility is

monitoring the learners, yielding feedback, and assisting the learners in realizing the areas that need improvement.

1.2. Teacher Assessment Literacy and Conceptualization

Xu and Brown (2016) introduced a framework conceptualizing assessment literacy, consisting of six components, including knowledge base, teacher conceptions of assessment, institutional and socio-cultural contexts, teacher assessment literacy in practice, teacher learning, and teacher identity (re)construction as assessors.

To conceptualize teacher assessment literacy for the present study, having the research question and the intention of the study in mind, and due to the direct relationship between Xu and Brown's (2016) suggested framework, the reconceptualization provided by them (which they refer to as assessment literacy in practice) is adopted in this study. The rationale for exploiting this framework is that firstly it considers the teachers from their knowledge-base to (re)constructing their identity as assessors, and secondly, the framework is in accordance with the purposes of the whole study which is related to the English language teachers' assessment literacy from planning to reflection on classroom-based assessments in Iran.

Certain features qualify assessment literate teachers. Stiggins (1991) and Koh (2011) list five standards of high-quality classroom assessment that teachers of the 21ˢᵗ century need to possess. These standards are 1) to start assessment with clear purposes; 2) to comprehend the significance of assessing various types of interrelated achievement targets; 3) to adopt appropriate methods for assessing different kinds of achievement goals; 4) to sample and collect the achievement of the students according to representative performance tasks; 5) to avoid assessment bias and misrepresentation that emerge from technical and practical problems.

Being assessment literate, according to Inbar-Lourie (2008), means "having the capacity to ask and answer critical questions about the purpose for assessment, about the fitness of the tool being used, about testing conditions, and about what is going to happen based on the results" (p.389).

Livingston and Hutchinson (2016) and Koh, Burke, Luke, Gong, and Tan (2018) have also discussed the characteristics of being assessment literate. In their opinion, educational professionals at all levels, particularly in the schools and classrooms, should know what students need to learn, how they learn, how teachers can promote learning, support it, and develop high-quality assessments, how to collect evidence concerning learning, and how to interpret that evidence and use it for improving learning.

Considering the components suggested by Xu and Brown (2016), this study sought to investigate the Iranian EFL school teachers' assessment literacy following the new curriculum shift and its direction toward assessment *for* learning. Also, as assessment *of* learning is the dominant paradigm in Iran's schools, its effect on teachers' classroom practices is investigated. Additionally, the impact of the educational system on teachers' classroom practices is discussed.

2. Literature Review

Teachers' assessment literacy has recently received considerable attention. White (2020) asserts that assessment literacy for teachers at all levels of classroom experience is recognized as a key competency. According to Hill (2017), the identification of the significance of developing teachers' assessment literacy is growing due to two reasons: (1) Because of the increment of using assessment for accountability, and (2) Because of the shift in emphasis in classroom-based assessment from AoL toward AfL. Scholars have discussed the importance of teacher training programs in developing teachers' assessment literacy.

In a considerable study, Tsagari and Vogt (2017) investigated the European EFL teachers' perception of language assessment literacy (LAL) levels in congruence with their expressed training needs. They had defined LAL as "the ability to design, develop and critically evaluate tests and other assessment procedures, as well as the ability to monitor, evaluate, grade and score assessments based on theoretical knowledge" (p. 377). The results of their study showed that LAL level was not enough for the intended assessment activities that were intended for execution in the professional field, teachers were unprepared for performing their assessment functions and the training provided in teacher training programs was insufficient.

Also, the study conducted by DeLuca and Klinger (2010), regarding the principles, policy, and practice concerning the assessment in education, discussed teacher training programs. Their study revealed as teacher training programs are deprived of mandatory assessment courses, a considerable number of teacher candidates compensate for this lack through practicum experiences. The drawback of this is that the knowledge and skills of associate teachers play a significant role in assessment quality.

The other study conducted by Djoub (2017) concluded that participants had not benefited from the essential training in educational assessment for maximizing its effectiveness. Their lack of assessment literacy was also realizable through their attitudes and practices in assessing their learners. Finally, she argues the need for training in assessment literacy as a fundamental part of teacher education programs.

In local studies, in 2018, Tavassoli and Farhady conducted a questionnaire-based study for investigating EFL teachers' needs in assessment knowledge. Their study revealed that a great ratio of the participants regarded the main topics of language assessment (testing language components/ skills, test development procedures, test characteristics, information about testing, and alternative assessment) as necessary or significant to be considered in language assessment courses.

Investigating language assessment literacy while focusing on the reformed policies in Iran, Firoozi, Razavipour, and Ahmadi (2019), reported the incompatibility of teachers' assessment perceptions with assessment preferences and goals assigned for the newly-developed curriculum. Particularly when one realizes that AfL is the intended practicum, but teachers' views are still AoL-oriented.

A constant concern in the literature was related to the teachers' assessment literacy and the challenges of training programs, regardless of their contexts and scopes. Most of the literature reviewed here showed that teacher training programs have not been effective in equipping teachers

with sufficient assessment literacy because they did not include adequate or appropriate input in their programs. This issue, in turn, leads to teachers' insufficient assessment literacy. Therefore, it's not surprising that some of the studies have reported that teachers' performance has not been satisfactory or their AL was rated poor. Additionally, what is missing in these studies is offering a solution for these drawbacks or the inappropriacy of teachers' assessment literacy, a gap that this project seeks to bridge to some extent at the end. In the meantime, Hildén and Fröjdendahl's (2018) study, conducted in Finland, achieved a different result: "by and large, the results suggest that the course objectives resonating the operative national curricula and topical scholar knowledge of student teachers' assessment literacy were attained fairly well" (p. 16).

2.1. Research Question

Considering the elements mentioned earlier, including Iranian EFL teachers' assessment literacy in conformity with the new curriculum, the educational system's impact on teachers' assessment practices, and the tensions that arisen, the following research question was formulated: RQ: How does the educational system impact teachers' classroom assessment knowledge and practices and create tensions in the classroom?

3. Context

3.1. The Educational System in Iran

In Iran, the Ministry of Education (MoE) is responsible for making educational decisions for the whole country. It initiated a reform in the educational system of Iran in 2010. Based on these reforms, the textbooks underwent radical changes and new textbooks were developed to enhance students' learning by concentrating on higher-order thinking processes. Concerning assessment, the schooling system in Iran has traditionally been extremely exam-oriented. The orientation toward summative assessment is predominant at the classroom level (Ghorbani, 2009).

The objectives that the new curriculum follows in teaching English in Iran, are based on the Common European Framework of Reference themes. They intend to help the learners to reach the A2 level at the end of junior high school and the B2 level at the end of senior high school (Foroozandeh & Forouzani 2015).

Based on the 2010 reform movement, the system of 6-3-3, which refers to the number of years students spend at school, was declared as the official agenda for schools. They spend six years at elementary school, three years at lower secondary school, and three years at upper secondary school. According to this movement, communicative language teaching was set as the default approach for teaching English in Iran, a shift that happened after 26 years of employing GTM (grammar-translation method) and reading methods of language pedagogy (Foroozandeh & Forouzani 2015). Students begin learning English in grade 7 when they are 12 years old. In grades 7 to 9, they study English 90 minutes a week, in grades 10 and 11, 3 hours, and in grade 12, 4 hours per week.

3.2. Classroom-Based Assessment Culture in Iran

To discuss the classroom assessment culture in Iran briefly, in the lower secondary school, students need to participate in at least two sets of exams. These two sets of teacher-made tests are administered at the end of the first and second term, referred to as final 1 and final 2 respectively. Besides, teachers themselves and sometimes schools, arrange some extra examinations before the finals to prepare the students for the main final exams.

The upper-secondary school is highly influenced by the university entrance examination (Konkoor). The attention of the students and the teachers at this level particularly in grade 12, is considerably shifted to that exam. The situation here is a clash between *AoL* and *AfL* for both levels, especially for the upper secondary school. The MoE expects teachers to vary their assessment and focus on assessment to support learning which means moving towards AfL. However, in reality, this has not happened yet. Additionally, it seems that a kind of misconception is taking place concerning conceptions that affect the whole direction of the teachers' activities. Explaining the assessment of the new books at lower secondary schools, Foroozandeh and Forouzani (2015) state that formative assessment or AfL, similar to summative assessment or AoL, is composed of 20 scores which contradicts the essence of AfL. AfL is not about scores, it's about the assistance, support, encouragement, and scaffolding that could be provided for the learners in the process of their advancement, monitoring them, and providing ongoing feedback (ARG, 2002; Fischer et al, 2011; Green, 2018; Mousavi, 2009). The tension between these two procedures is challenging for the teachers. On one hand, they are required to settle between providing effective teaching that helps learners achieve the instructional aims including preparing learners for communicative purposes, creating self-confidence and motivation in the learners, and developing the four skills at the intermediate level (Alavi-moghaddam et al., 2018). On the other hand, they have to prepare the students for the final exams and for the Konkoor, exams which are different.

Concerning providing effective feedback, based on Boyd, Green, Hopfenbeck, and Stobart (2019), seven elements exist in giving feedback. They believe that effective feedback: *is specific and clear*: the learner needs to be told where exactly they need to practice and what to improve; *is well-timed:* in deciding about the appropriate time for giving feedback. Learners' levels, the task nature, and feedback type are influential in this regard; *is clearly linked to the learning intention:* if the intention is grammar, then the students should receive feedback on the grammar and not handwriting or spelling, for instance; *focuses on the task rather than the learner:* what matters in this element is the "how" not only the "what" of the feedback. Students' personalities affect the quality of the feedback. If we give a shy student a sharp and disappointing remark directly, it might impede their improvement or they might lose their interest which are totally against the purpose of giving feedback; *gives prompts at the right levels on how to move learning forward:* a significant skill the teachers need in providing feedback is adjusting the type of the feedback with their learners' levels; *offers strategies rather than solutions:* to make the students more independent in their learning process, they will need a type of feedback which encourages them to reflect on their own performance; and *challenges learners, requires action, and is achievable*: the feedback should

be challenging. To achieve this purpose, it should be specific and tell the learner what to do to achieve the desired outcome.

4. Method

4.1. Participants

To get in touch with the interviewees, the secretariat of the EFL teachers' national group was contacted, and they provided the teachers' contact information. At the next step, teachers were sent invitations for the interviews. The interviewees were 15 upper-secondary school headteachers from different cities in Iran. The total number of headteachers equaled 36 teachers. Twenty teachers were invited for the interviews. Some of them were not willing to participate or did not respond to the calls. The logic behind interviewing headteachers is that from among their activities, they usually review the exams developed by their colleagues in their area to discuss them in their meetings and improve their qualities. Consequently, they could delineate a general picture of the teacher-made tests. The background information of the participants is provided in the following table.

Table 1

Participants' Background Information

Teachers	Gender	Experience (years)	Degree
Teacher 1	M	22	MA
Teacher 2	F	23	MA
Teacher 3	M	31	MA (Linguistics)
Teacher 4	F	25	MA
Teacher 5	M	16	Ph.D. Candidate (TEFL)
Teacher 6	M	25	Ph.D. Candidate (Linguistics)
Teacher 7	F	33	MA
Teacher 8	F	29	Ph.D. (Psychology)
Teacher 9	M	24	MA
Teacher 10	M	27	MA
Teacher 11	F	25	Ph.D.
Teacher 12	F	22	Ph.D. Candidate (TEFL)
Teacher 13	M	22	MA
Teacher 14	M	26	Ph.D. (TEFL)
Teacher 15	F	26	MA

4.2. Instruments and Data Collection

The data was gathered through interviews accompanied by questionnaires. Cohen, Manion, and Morrison (2011) maintain that interviews allow for great depth in comparison to other research

methods. Both of the instruments were utilized to assist in comprehending teachers' needs, challenges, tensions, and attitudes from their viewpoints, as deeply and thoroughly as possible.

An adjusted version of Vogt and Tsagari's (2014) questionnaire (Table 2) was employed to explore the teachers' needs in assessment literacy. The researcher adjusted the questionnaire by eliminating some items that did not fit the context of the study and applying some modifications to the items. The internal consistency reliability of the original questionnaire composed by Vogt and Tsagari (2014) was computed using Cronbach's alpha, which ranged from $\alpha = .80$ to $\alpha = .93$, indicating a high level of internal consistency. Also, the internal consistency of the adjusted version was calculated and it was $\alpha = .86$ which was consistent with the original questionnaire.

The timetables of the interviews were finalized based on the interviewees' convenience and preferences. Since the participants were from different parts of the country, the interviews were conducted through audio phone calls. The duration of the interviews ranged from 30 to 45 minutes. The interviews and the questionnaires are discussed respectively in what follows.

4.2.1. Interviews. The interviews were semi-structured as the guiding items (Appendix 1) allowed addition or variations and did not limit the depth and breadth of the respondent's thoughts (Dörnyei, 2007). The interviews were conducted in Persian, Iran's standard national language for the sake of convenience. To ensure that the transcripts remained maximally faithful to the audio content, the interviews were then transcribed and translated by the researcher, as he holds a university degree in translation and has some years of experience in transcribing and translating English to Persian and vice versa materials.

Each interview began by providing a short orientation as to how the interview would proceed. Participants were informed that the interviews are being recorded. The interviewees were free to discuss the items as long as no major regression from the main theme occurred. In that case, the researcher tried to redirect the conversation by raising a related question. The researcher tried not to interrupt to keep the natural flow of the discussion. Whenever there were some ambiguities, the researcher tried to provide a summary of the main point to ascertain that comprehension has happened appropriately.

4.2.2. Analysis of the interviews. The transcripts of the interviews were reviewed and coded for content analysis, considering the main research question and the themes of the interview items. Content analysis was used because it makes qualitative inferences possible by analyzing the meaning and semantic relationship of words and concepts (Mousavi, 2009). The interviews transcripts were anonymized by replacing the names of the participants with numbers. The codes of the analysis were generated through grounded coding and the researcher and an expert in language testing and assessment coded the interview transcripts independently and discussed discordant codes, coming to a joint decision concerning each of such cases.

4.2.3. English Language Teachers' (ELTs') Questionnaire. An adjusted version of Vogt and Tsagari's (2014) questionnaire was employed for the present study. It was justified to meet the needs of the context of the present study. Those items that were related to the context of the study and matched the requirements of the study were used from the aforementioned questionnaire. The questionnaire included 14 items (Table 2), and teachers were asked to clarify if they were

trained, need training, or were not trained academically but learned on the job. The items are presented in Table 2.

Table 2
Questionnaire Items

A: Classroom-focused language testing and assessment
[1] Preparing [effective] classroom tests (which include all the components)
[2] Using ready-made tests from textbook packages or other sources
[3] Giving feedback to students based on information from tests/assessment
[4] Using self- or peer-assessment
[5] Using an informal, continuous, non-test type of assessment
B: Purposes of testing & assessment
[6] Giving grades (assigning grades to sub-items or skills appropriately)
[7] Finding out what needs to be taught/learned
C: Content and concept of language testing and assessment
[8] Testing/Assessing: Receptive skills (reading/listening)
[9] Testing/Assessing: Productive skills (speaking/writing)
[10] Testing/Assessing: Microlinguistic aspects (grammar/vocabulary)
[11] Testing/Assessing: Integrated language skills
[12] Establishing the reliability of tests/assessment
[13] Establishing the validity of tests/assessment
[14] Using statistics to analyze the quality of tests/assessment

5. Results

5.1. Interviews

Based on the analysis of the interview data, some themes were more prominent regarding their function in the classroom and their impact on both teachers' and learners' performances. The main themes and their relationship with Xu and Brown's (2016) conceptual framework, where relevant, are presented in detail in the following subsections.

1. Focusing on cheating rather than helping learners to develop

The first item in the interviews was related to the reliability of online assessment during the pandemic (Q1). The majority of the teachers asserted that concerning reliability, the exams weren't reliable at all due to socio-cultural reasons and infrastructural facilities. Concerning the former, teachers were worried because of the students' widespread cheating. For instance, Teacher 8 stated, "from the socio-cultural perspective, it was probably satisfying to two-three percent only." Cheating was one of the teachers' main preoccupations, if not the most prominent one. The teachers, therefore, mainly discussed challenges in assessment regarding the assessment of learning outcomes rather than the learning process and focusing on objectivity and reliability of assessment. From the infrastructural point of view, the students had some technical problems that could undoubtedly affect their performances.

2. Navigation between AoL and AfL

The navigation and struggle in grade 12 are related to the pre-determined exams, the format of which is dictated by the MoE. The other important point was related to the main sources that teachers used for developing their assessments (Q6). They included: teachers' self-designed exams, colleagues designed exams, commercially prepared materials, internet websites, and school materials. Teachers had various views on the sources. Regarding colleagues' items some teachers posed contemplative points, for instance, Teacher 3 believed that "the negative point is that they might not be in line with or at the level of what you have taught" and some teachers believed that colleagues' exams suffer from grammatical drawbacks, such as Teacher 5 who asserted "about 80 percent of these exams suffer from grammatical mistakes". Concerning using commercially prepared materials, Teacher 9 expressed that "if I use the commercially prepared books, it means that I'm not capable of developing my exam". Teachers' views are following Hill's (2017) attitudes: "sourcing assessment tasks online, from textbooks, or other teachers, for example, raises issues of task quality and fitness to purpose". Also, considering the research question of the present study, Teacher 1 highlighted the tension regarding classroom assessment "Teachers should consider the students' activities during the whole semester; but don't underestimate their assessment or make it too easy; rely on formative assessment rather than summative assessment."

3. Top-down decisions

The MoE authorities are the decision-makers regarding everything related to the instruction, including course content and level, the time allowed per course, and the assessment policy. The teachers are performers of their decisions without playing a significant or active role. As teachers 3 and 5 commented, the decisions are made by the authorities which (regarding assessments) lead them and the students to care about the scores only and not learning. Teacher 9 added that "we have a lot of groups for the teachers and the headteachers, however, they have no role in (a) deciding about the workshops that are planned for teachers, (b) the workshops' instructors, (c) their content, and (d) on teachers' qualifications".

This theme is consistent with "micro-and macro-contexts as the boundaries for teachers' assessment literacy in practice component", which argues that despite the decisive role of teacher conceptions of assessment, in-service teachers are instructors only and their decision-making ability is highly restricted.

4. Teachers' agency

Teachers perceive the final nationwide exam of grade 12 and the Konkoor as impediments to their classroom practices. As Teacher 9 asserted "I have tried various forms of exams, but unfortunately and unfortunately, our final goal should be focusing on the Konkoor and we do so." Teacher 1 stated that "Actually, I try to make them [the exams] various, but the truth is that we have to follow an unwritten rule, particularly for grade 12. We have to prepare the students for that exam." Supportively, Teacher 6 added that "Due to the weight of Konkoor exam and the attention that students pay to it, I usually give only one essay type exam. However, every two weeks I give a multiple-choice exam based on the Konkoor exam." Teacher 3 explained that the educational system is the decision-maker regarding the exams and the teachers have no role in the exams.

Teacher 4 emphasized, "well, I use the format of the final exams as the foundation of my exams because our education is result-based." Similarly, Teacher 5 added that "our education system is mainly score-based. The education system leads me and the students to care about the score only and not learning." Attitudes expressed by these teachers strongly manifest that the educational system needs to revise its policy towards providing further freedom for the teachers concerning preparing or at least selecting their classroom materials, and also respecting them in their assessment decisions. This theme is also in line with "micro-and macro-contexts as the boundaries for teachers' assessment literacy in practice component", discussed in the previous theme.

5. *Adjusting testing to teaching*

Teachers believed that when they develop their exams and adjust their assessment to what they have taught, that would make their assessment fairer, they could evaluate their students more precisely, and the results would be more reliable (Q3). Teacher 2 stated, "The only thing that I do and I think is effective, is that I design my questions and do not copy them,". The most prominent and outstanding concept that was noticeable among the teachers was the employment of and appealing to evaluation and assessment for creating a link between teaching and testing, a linkage that Vogt and Tsagari (2014), consider as one of the elements of classroom-focused assessments. Among the participants, 13 teachers believed that keeping the level of the exams in agreement with the teaching level could help learners achieve more acceptable results and it gives the teachers an overall picture of their learners' progress.

Considering the "knowledge base" component suggested in Xu and Brown's (2016) framework, this theme is related to the sub-component of "knowledge of assessment purposes, content, and methods" in which the content of the exam is supposed to match what has been taught.

6. *Assessment fairness*

As the interviewed teachers were headteachers, they were also asked about advising junior colleagues about making their assessment fair (Q7). By fair assessment, I mean "achievement evaluation based on the course materials". The point that teachers highlighted repeatedly was "to test what we have taught" which is discussed above. The other recommendation was perhaps relying on several assessments instead of just one exam, which was clear in Teacher 4 explanations:

> "I advise less-experienced colleagues not to rely on the final exam solely, take students' performance during the term into consideration, distribute the score to various activities not only the formal exams scores. I also recommend them to print their exam, revise it if needed and answer the questions themselves. They may find some minor mistakes in it"

Considering the students' levels, their performances during the semester, and considering their active participation in classroom activities, were the other factors that were highlighted in the teachers' responses. Considering the students' level and adjusting classroom assessment to their level can give the teachers a clear picture of where the students are and how the teachers can lead them where they need to go with their learning. Observing the learners' performance in the

classroom and considering their active participation in the process of learning are also helpful in assessing them more fairly.

7. Power of high-stakes exams

The interviews confirmed that teachers find the Konkoor exam considerably interfering with their classroom practices. It was evident in Teacher 9 comment who expressed that "I have to acknowledge that we have to have an eye on the final exam and the Konkoor." Teacher 6 asserted that "due to the weight of Konkoor exam and the attention that students pay to it, I usually give only one essay type exam at grade 12. However, every two weeks I give a multiple-choice exam based on Konkoor" which could be considered a negative washback. Teacher 11 critically asserted that:

> "As we have to prepare the students for the final exams, we move in that direction. We are resolved in this unhealthy policy even if we don't like or agree with it. Consequently, in the first term, we act based on our discernment, and in the second term, we try to prepare them for the final exams".

In this theme, "teacher identity (re)construction as assessors" is the related component. This component discusses teachers' traditional role which means that teachers are "involved in drilling and repeatedly preparing students for success on an externally produced examination" (Xu & Brown, 2016) and their role under formative assessment policies based on which "teachers assess students to make pedagogical decisions, and these decisions are sometimes used as part of formal certification processes." (Xu & Brown, 2016). The traditional role is more compatible with Iran, the context of this study.

8. Feedback

Teachers were aware of the importance of giving feedback, but in practice, it looked that feedback did not receive much emphasis. The findings of the present study revealed that feedback in the participating teachers' practices was restricted to two elements: (1) specific and clear feedback and (2) well-timed feedback. As an example of the former, Teacher 15 stated "I provide comments on their papers … and encourage discovery learning." and as an example of the latter, Teacher 1 described it as "I leave comments on their papers and also explain after giving the corrected papers back to the students."

"Knowledge of feedback, teacher learning, and teacher assessment literacy in practice" are the components of the framework matching this theme. The teacher learning component considers teachers' participation in activities concerned with assessment (e.g., moderation, rubric development, assessment task design) and real virtual networking. "Teacher assessment literacy in practice" is also connected to this theme because teachers' assessment-based decision making and subsequently, the action they take based on them, determine if teachers "treat assessment as a quality assurance mechanism or learning-oriented tool." (Xu & Brown, 2016).

5.2. ELTs' Questionnaire

The main purpose of the questionnaire was to find out about teachers' needs in the process of assessment and the training that supports their practices. Descriptive reports of the questionnaire results and teachers' prioritization of their training needs are presented in the following table.

Table 3
Teachers' Prioritization of their assessment needs

Item	Area	Frequency (selected by)
[14] Using statistics to analyze the quality of tests/assessment	C	10 teachers
[4] Using self- or peer-assessment	A	8
[7] Finding out what needs to be taught/learned	B	8
[8] Testing/Assessing: Receptive skills (reading/listening)	C	8
[9] Testing/Assessing: Productive skills (speaking/writing)	C	8
[11] Testing/Assessing: Integrated language skills	C	8
[12] Establishing the reliability of tests/assessment	C	8
[13] Establishing the validity of tests/assessment	C	8
[1] Preparing [effective] classroom tests (which combine all the components)	A	6
[5] Using an informal, continuous, non-test type of assessment	A	6
[6] Giving grades (assigning grades to sub-items or skills appropriately)	B	6
[3] Giving feedback to students based on information from tests/assessment	A	6
[2] Using ready-made tests from textbook packages or other sources	A	5
[10] Testing/Assessing: Microlinguistic aspects (grammar/vocabulary)	C	4

Considering the frequency of the items prioritized by the teachers as their needs, which was the purpose of this part of the study, it could be concluded that "C: content and concept of language testing and assessment" received the highest amount of attention. After that come "A: classroom-focused language testing and assessment" and "B: purpose of testing" respectively. In other words, teachers' needs lie in the practicalities of assessment.

6. Discussion and Implications

The main purpose of the current study was to survey the impact of the educational system on teachers' classroom practices, their classroom assessment knowledge, and the tensions that arise from these factors. The study showed that the teachers were not satisfied with the reliability of the exams during the online period. Socio-cultural factors and the available facilities were important factors in this regard. The role that the Ministry of Education was playing in education was also effective in the teachers' classroom performance and practices. Teachers' felt that their path is somehow predetermined and they cannot imply many changes in their teaching process.

Teachers also navigate between assessment for and of learning. On the one hand, the educational system requires them to apply AfL, on the other hand, it has not provided the required training. It causes some tension for the teachers. Top-down decisions, taken mainly by the MoE, and teachers' passive role in the decisions, lead teachers and the students to prioritize scores over learning. The educational system is also the decision-maker regarding the assessment and the teachers have a little role in the exams. Attitudes expressed by these teachers strongly target teachers' agency and manifest the educational system's need for revising its policy towards providing further freedom for the teachers concerning assessment and selection of classroom materials.

Consistent with Tsagari and Vogt, (2017), Djoub (2017), and Tavassoli and Farhady (2018), the findings of the present study revealed the need for a professional development program in assessment literacy, particularly regarding the purposeful use of AfL and AoL. Concerning the use of AfL and AoL, the interviews finding revealed that AoL is highly dominant and the AfL is not considered practically. Teachers do not consider it in their classroom assessments. Some teachers referred to it, however, they regarded it the same way that Foroozandeh and Forouzani (2015) had discussed it, i.e., setting scores for it, while scientifically, AfL is not engaged with scores, it is for promoting learning. It indicates that a serious process-based movement toward AfL, accompanied by patience, is needed if the education system intends to apply AfL in the classroom. As Livingston and Hutchinson (2016) assert, the process of transformation from summative only to a purposeful mixture of assessments, with the ability of continuous renewal over time and based on new needs, should be realized as a process rather than an event.

Moreover, teachers' passive agency and their ignored role in decision-making are not desired by teachers and they find them annoying. This finding corresponds with the results of Tsagari and Vogt's (2017) study which manifested that foreign language teachers think about their positionality or agency in language assessment. Policy and decision-makers are addressed in this regard. As "teachers have the capacity to be agents of change where school policies related to testing and grading are concerned" (Clark-Gareca, 2019, p. 56), they should be granted freedom in their classroom assessment practices.

Various reasons could be counted for teachers' lack of assessment literacy. One of the main reasons is related to teacher education programs and their failure in providing adequate training in assessment (Tsagari & Vogt, 2017) and their effectiveness (White, 2019). The other one, according to Djoub (2017), could be traced back to the restricted vision of the focus of assessment. Finally, based on the results of Firoozi, Razavipour, and Ahmadi's (2019) study, to do effective assessments, teachers are required to possess appropriate knowledge about the curriculum, content, and the subject matter, regardless of the educational context.

The findings of the present study could prove helpful to three stakeholders, namely teachers, teacher trainers, and policymakers. Because a considerable amount of assessment occurs in classrooms, teachers' role as an important group of stakeholders is underlined which may cause the teachers to be seen as the real protagonists in language testing and assessment (Tsagari, 2021). Teachers can benefit from the findings by developing their knowledge and assessment literacy, especially concerning assessment for learning and appealing to it for ameliorating their learners'

achievements. Feedback, as the current study revealed, is an important factor that can assist teachers in helping their learners. However, in comparison with the seven elements proposed by Boyd et al. (2019), five elements are missing in the practices of the participating teachers. Other researchers can also investigate how to develop teachers' knowledge in giving effective feedback, considering the elements proposed by Boyd, et al. (2019).

Teacher trainers can play an essential role in giving direction to the student teachers at teacher training programs and help in implementing AfL in their practices. Contrary to teachers who have little control over their materials in Iran, teacher trainers at universities have freedom in compiling or selecting their materials. Therefore, they can play a considerable role in making this transformation towards needed formative assessments possible. As a subject for further research, it sounds interesting to investigate teacher trainers' willingness and attitudes towards this transformation and their preparedness for applying it.

Finally, the policymakers' role is definitely of great importance, as they are called for language assessment literacy promotion (Kremmel & Hrding, 2020). They can revise what they have done so far and what results have been obtained. Based on the findings of this study, their decisions have affected the teachers' practices, motivation, and agency. Most of the teachers revealed their dissatisfaction specifically with their lack of agency. The whole decision of directing teaching toward communicative language teaching is, without a doubt, positive, however, it doesn't occur if the teachers are not trained sufficiently and the infrastructures, including the facilities (such as audiovisual aids) required for implementing it in the classrooms, are not prepared for that.

7. Conclusion and Limitations

The present study aimed at understanding EFL teachers' assessment literacy in Iran, teachers' needs, and exploring their attitudes towards assessment *for* learning based on the information given by the headteachers as informants. In addition, the dominance of the AoL effect on teachers' classroom practices was sought to be investigated through interviews and questionnaires.

The needs prioritization of the questionnaire items by the teachers revealed that items that dealt with assessing receptive, productive, Microlinguistic aspects, integrated language skills, reliability, validity, and the usage of statistics for analyzing the quality of the evaluations and communicating the results were the most challenging ones. After that come "classroom-focused language testing and assessment" and "purpose of testing" respectively. In other words, teachers' needs lie in the practicalities of the evaluation. Although the number of the participants was not high enough to make generalizations possible, nor did the study mean it at this level, the purpose of the present study, i.e., realizing teachers' needs for developing a questionnaire, was achieved and the findings of the present study will pave the way for that purpose.

Analyzing the interviews and coding them, manifested that eight themes were fundamental regarding assessment and were affecting the teachers' classroom practices which are 1) Focusing on cheating rather than helping learners to develop, 2) Navigation between AoL and AfL

preparation approaches, 3) Top-down decisions, 4) Teachers' agency, 5) Adjusting testing to teaching, 6) Assessment fairness, 7) Power of high stakes exams, and 8) Feedback.

Finally, it should be emphasized that movement towards assessment culture based on formative assessment is not the responsibility of the teachers alone. The decision-making authorities, school administration, the students, and their parents need to be open-minded regarding this change and be willing to cooperate in applying it (Verhoeven & Devos, 2007). Changing teachers' classroom assessment culture is highly challenging and requires the investment of an enormous amount of resources (Hazim Jawad, 2020). Individual and lonely efforts implemented in an uncomprehensive way would not yield long-lasting and reliable changes. "Rather, this requires professional, cultural, and system changes to be conceptualized as multiple cogs and wheels that all have to be operational and interacting continuously to be effective." (Livingston & Hutchinson, 2016, p. 16).

Resembling many other studies, my study is also affected by some limitations that influence the generalizability of the results. A larger number of participants is needed to improve the validity of the findings. In addition, analyzing the real assessments performed in the classrooms can also delineate a clearer picture of classroom assessments in practice that can assist in obtaining extra sources of data for gaining more detailed insights concerning the nature and effect of language assessment practices on language learning.

References
Alavi-moghaddam, B., Kheir-abadai, R., Rahimi, M., & Davari, H. (2018). *Vision 3: English for schools: Teacher's guide.* Chap va Nashr Iran: Tehran.

Bennet, J. (2017). Assessment for learning vs. assessment of learning. Retrieved from https://www.pearsonassessments.com/professional-assessments/blog webinars/blog/2017/12/assessment-for-learning-vs--assessment-of-learning.html

Boyd. E, Green A, Hopfenbeck. TN, & Stobart. G (2019) 'Effective feedback: the key to successful assessment for learning', *ELT position papers*: Oxford University Press.

Clark-Gareca, B. (2019). Assessing pre-service teachers' assessment literacy: Building a

foundation. In E. White, & T. Delaney (Eds), *Handbook of Research on Assessment Literacy and Teacher-Made Testing in the Language Classroom* (1st ed., pp. 42-57). IGI Global

Cohen, L., Manion, L., & Morrison, K. (2011). *Research Methods in Education* (7th ed.). London: Routledge.

Coombe, C., Vafadar, H., & Mohebbi, H. (2020). Language assessment literacy: What do we need to learn, unlearn, and relearn? *Language Testing in Asia*, *10*(3), 1-16, https://doi.org/10.1186/s40468-020-00101-6

Deluca, C., LaPointe-McEwan, D., & Luhanga U. (2016). Approaches to classroom assessment inventory: A new instrument to support teacher assessment literacy. *Educational Assessment*, *21*(4), 248-266. https://doi.org/10.1080/10627197.2016.1236677

Deluca, C., & Klinger, D. (2010). Assessment literacy development: Identifying gaps in teacher candidates' learning. *Assessment in Education: Principles, Policy & Practice*, *17*(4), 419-438, DOI: 10.1080/0969594X.2010.516643

Djoub, Z. (2017). Assessment Literacy: Beyond teacher practice. In R. Al-Mahrooqi et al. (Eds.), *Revisiting EFL Assessment, Second Language Learning and Teaching*, DOI 10.1007/978-3-319-32601-6_2

Dörnyei, Z. (2007). Research methods in applied linguistics. Oxford, UK: Oxford University Press.

Firoozi, T., Razavipour, K., & Ahmadi, A. (2019). The language assessment literacy needs of Iranian EFL teachers with a focus on reformed assessment policies. *Language Testing in Asia*, 9(2). https://doi.org/10.1186/s40468-019-0078-7

Fischer, J., Chouissa, C., Dugovičová, S. & Virkkunen-Fullenwider, A. (2011). *Guidelines for task-based university language testing.* European Centre for Modern Languages. Council of Europe Publishing

Foroozandeh, E., & Forouzani, M. (2015). Developing school English materials for the new Iranian educational system. In C. Kennedy (Ed). *English language teaching in the Islamic Republic of Iran: Innovations, trends and challenges* (pp. 59–73). British Council

Ghorbani, M.R. (2009). ELT in Iranian high schools in Iran, Malaysia and Japan: Reflections on how tests influence use of prescribed textbooks. *Reflections on English Language Teaching*, *8*(2), 131-139

Green, A. (2016). Testing four skills in Japan. *British Council New Directions in Language Assessment: JASELE Journal, (Special Edition),* 135-144.

Green, A. (2018). Assessment for learning in language education. *Iranian Journal of Language Teaching Research*, 6(3), 9-18. DOI: 10.30466/ ijltr.2018.120598

Hazim Jawad, A. (2020). Examination of Iraqi EFL teachers' attitudes, intentions, and practices regarding formative assessment. *International Journal of Language Testing*, *10*(2), 145-166.

Hildén, R., & Fröjdendahl, B. (2018). The dawn of assessment literacy – exploring the conceptions of Finnish student teachers in foreign languages. *Apples – Journal of Applied Language Studies*, 12(1), 1–24. https://doi.org/10.17011/apples/urn.201802201542

Hill, K. (2017). Understanding classroom-based assessment practices: A precondition for teacher assessment literacy. *Papers in Language Testing and Assessment*, 6(1), 1–17.

Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*, *25*(3), 385–402. https://doi.org/10.1177/0265532208090158

Inbar-Lourie O. (2017) Language Assessment Literacy. In: Shohamy E., Or I., May S. (Eds) *Language Testing and Assessment. Encyclopedia of Language and Education* (3rd ed.). Springer, Cham. https://doi.org/10.1007/978-3-319-02261-1_19

Koh, K.H. (2011). Improving teachers' assessment literacy through professional development. *Teaching Education*, 22(3), 255-276, DOI: 10.1080/10476210.2011.593164

Koh, K., Burke, L. E. C.-A., Luke, A., Gong, W., & Tan, C. (2018). Developing the assessment literacy of teachers in Chinese language classrooms: A focus on assessment task design. *Language Teaching Research*, *22*(3), 264–288. https://doi.org/10.1177/1362168816684366

Kremmel. B., & Harding. K. (2020) Towards a comprehensive, empirical model of language assessment literacy across stakeholder groups: Developing the language assessment literacy survey. *Language Assessment Quarterly*, 17:1, 100-120, DOI: 10.1080/15434303.2019.1674855

Livingston. K., & Hutchinson. C. (2016): Developing teachers' capacities in assessment through career-long professional learning. *Assessment in Education: Principles, Policy & Practice*, DOI: 10.1080/0969594X.2016.1223016

Luxia, Q. (2007). Is testing an efficient agent for pedagogical change? Examining the intended washback of the writing task in a high-stakes English test in China, *Assessment in Education: Principles, Policy & Practice*, 14:1, 51-74, DOI:10.1080/09695940701272856

Mousavi, S.A. (2009). *An encyclopedic dictionary of language testing*. Tehran: Rahnama Press

Razavipour, K., & Rezagah, K. (2018). Language assessment in the new English curriculum in Iran: Managerial, institutional, and professional barriers. *Language Testing in Asia*, *8*(1), 1–18.

Stiggins, R. (1991). Relevant Classroom Assessment Training for Teachers. *Educational Measurement*, 10(1), 7–12.

Tavassoli, K., & Farhady, H. (2018). Assessment knowledge needs of EFL teachers. *Teaching English Language*, 12(2), 45-65.

Tsagari, D. (2021). Gauging the assessment literacy levels of English language teachers in Norway. *The European Journal of Applied Linguistics and TEFL*. 10(1), 161-191.

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*

Van der Kleij, F.B., Cumming. J.J., & Looney, A. (2018) Policy expectations and support for teacher formative assessment in Australian education reform. *Assessment in Education: Principles, Policy & Practice*, 25(6), 620-637. DOI: 10.1080/0969594X.2017.1374924

Verhoeven. J.C & Devos. G. (2005) School assessment policy and practice in Belgian secondary education with specific reference to vocational education and training. *Assessment in Education: Principles, Policy & Practice*, 12:3, 255-274, DOI:10.1080/0969594050033723

Vogt, K. & Tsagari, D. (2014) Assessment literacy of foreign language teachers: Findings of a European study. *Language Assessment Quarterly*, 11(4), 374-402, DOI:10.1080/15434303.2014.960046

White, E. (2019). Developing assessment literacy through assessing classroom tests: Instruments and procedures. In E. White, & T. Delaney (Eds), *Handbook of Research on Assessment Literacy and Teacher-Made Testing in the Language Classroom* (1st ed., pp. 16-41). IGI Global

White, E. (2020). Are you classroom assessment literate? Self-assessing your assessment competence. In E. White & T. Bucher (Eds), *Classroom Assessment for Language Teaching: Challenges, Choices, and Consequences* (1st ed., pp. 326-338). Cambridge Scholars.

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

Appendix 1: Interview Guiding Items

1. How was your year? How reliable was online evaluation in your assessment?

2. How do you know what your students know or have learned?

3. How do you think a teacher can tailor his/her assessment to the learners' existing knowledge and how do you usually do it?

4. What do/don't you tell your students about assessments?

5. Based on your experience, could encouraging the learners to discuss the basis of their responses affect their learning? (If yes, how? if no, why?)

6. Where do your formal assessment activities usually come from? (e.g., school textbooks; self-designed; other teachers; commercially-prepared materials; a mixture of them)

7. How would you advise a junior colleague about making an assessment fair? (Can you give an example?) What challenges might be there?

8. How do you think assessment could help in diagnosing the learners' needs in planning teaching?

9. What might a good/bad test result tell you?

10. How do you give feedback mainly? (Written, oral, individually, general (to the whole class))

11. Based on your own experience, how well do students use their teacher's feedback for their improvement?

12. In your opinion, how can I (as a novice teacher) understand if the learners have understood the feedback and engaged with it?

13. How could it affect the learners' performance if they understand the purpose of the assessment?