Turkish EFL instructors' in-class language assessment literacy: perceptions and practices

Serpil Öz[1]

Bahçeşehir University, İstanbul-Turkey

Derin Atay[2]

Bahçeşehir University, İstanbul-Turkey

Abstract

The current study was conducted with the aim of understanding Turkish EFL instructors' perceptions towards in-class language assessment and its reflection in their classroom practice. The purposes of this research are: (1) to find out Turkish EFL instructors' perceptions of the term 'in-class language assessment', (2) to reveal their reflections of their in-class assessment knowledge to their own practice and (3) to find a relationship between experience and perception of in-class assessment. In this study, data was collected from twelve instructors, 8 (female) and 4 (male) prep school teachers, varying in experience between 1-15 years, working in English Preparatory Program of a Turkish university. For the study, semi-structured interview was applied in order to collect data. The interviews were recorded and transcribed through reoccurring content and themes. The findings of the study show that, although most of the teachers were familiar with basic classroom assessment, when it comes to classroom practice, there is an imbalance between assessment literacy and classroom reflection. It also revealed that there is not much relationship between the experience and assessment perception. The findings have some implications for the teachers, continuous professional units, school principals and teacher educators in the area of testing and have some recommendations for the future studies.

Keywords: Assessment Literacy, In-class Language Assessment, In-class Assessment Practice

---

[1] Instructor of English, Bahçeşehir University. **Email:** serpil92oz@gmail.com

[2] Prof., Dr., Bahçeşehir University, Foreign Language Education Department. **Email:** derin.atay@es.bau.edu.tr

Introduction

At the very beginning, it is vital to differentiate testing, assessment and evaluation. Assessment is using all types of measurements in order to see the progress whereas testing is the subcategory of assessment and it is more of a serious term which includes methodical and standardized measurement (Coombe, Folse, & Hubley, 2007, p. xv). Then, it is a need to analyze assessment in a larger perspective; it is not only pen-paper type of measurement. Thomas, Allman, & Beech (2004) define assessment as collecting data which helps to use in decision-making process (p. 2). Also Susuwele-Banda (2005) supports the very same definition and describes as "The process of collecting information purposefully using different methods/strategies and tools for the purposes of informing decision" (p. 11). On the other hand, assessment and evaluation are related terms; evaluation is decision-making about the related program/course but assessment is measuring that program/course (Kondal, 2015, p. 241).

Assessment has many benefits for education. It reflects teacher success and shows what goes right, what goes wrong or what parts need some change and improvement. With the help of assessment, teachers are able to understand and track the flow of the teaching and make necessary changes in order to improve and enhance learning of the students. Coombe et al. (2007) define assessment as "*Assessment* refers to a variety of ways of collecting information on a learner's language ability or achievement" (p. xv). In that way, Llosas (2013) claims that it gives a common framework for the teachers to "improve the quality of education for all students by developing rigorous standards and aligning instruction and assessment" (as cited in Jannati, 2015, p. 27). Also Fleming (2007) adds another advantage of assessment as "…determines what kind of feedback is needed to improve pupils' learning" (p. 2). Rea-Dickins (2004) states the importance of assessment for the teacher in every step as "In making decisions about lesson content and sequencing, about materials, learning tasks and so forth teachers have to determine the strengths and weaknesses of alternatives available to them." So assessment has an essential role in teacher' decision making process. It goes with the teacher nearly in every step.

Considering education, assessment is an undeniable need; without assessment, it is nearly impossible to determine how the progress goes on, whether students are successful or whether curriculum works well or not. As Rust (2002) mentions the importance of assessment as "...it determines much of the work students undertake (possibly all in the case of the most strategic student), affects their approach to learning..." (p. 1). While assessment is that much important, according to the studies, there is some inefficiency when it comes to the real classroom practice.

Little research is conducted on the reasons of teachers' inefficiency when it comes to the in-class language assessment itself. Limited studies attribute to the insufficiency of pre-service education (Gok, Erdogan, Altinkaynak, & Erdogan, 2012; Shepard & Sheppard, 2000; Kahl, Hofman, & Bryant, 2012) claiming that pre-service education is not enough and does not reflect real classroom practice. On the other hand, few studies claim that assessment insufficiency occurs due to teacher being 'assessment illiterate' (Jannati, 2015; Susuwele-Banda, 2005; Mertler, 2003), while the others claim that teachers have misperceptions

(Montee, Bach, Donovan, & Thompson, 2013) or assessments being scary itself since there is much to know about it (Salvia, Ysseldyke, & Witmer, 2012, p. 17). Rogers (1991) state that teachers need more training due to their lack of knowledge in terms of assessment (p. 181). Also teachers report that they are not fully equipped with the knowledge of assessing the students (Plake, 1993). So, teacher's applying effective assessment skills reflects better classroom outcomes. It is a need to develop teachers' judgmental skills and improve their assessment literacy in order to function better in educational settings. In this way, this study was conducted and the purpose of this study was to identify what Turkish EFL instructors' viewpoints about language assessment as an educator and how they reflect their knowledge of in-class language assessment to their practices and if there is any, to reveal the relationship between experience and perception of in-class language assessment.

Theoretical Background

Language Assessment

Assessing a language is multidimensional job for language teachers since it requires assessing four main basic skills of English: listening, speaking, reading, writing. Each skill has different assessment tools, approaches and techniques in an educational setting. So it is the teacher's job to decide when and how to use those tools and techniques. All those things that mentioned above are complicated and require strong decision-making process. Meanly, teacher is in the center taking care of the assessment progress and tries to provide success for the students and the program as well. Allman & Beech (2004) claim that if the learners are successful, that is mainly teacher's wise choices of instructional designs, curriculum and assessment; it is like a mirror effect (as cited in Yamtim & Wongwanich, 2014, p. 2999).

In order to make the learning more efficient, it is necessary to observe teachers' usage of assessment to make improvement in necessary parts. Gok et al. (2012) support the idea that "Whether measurement and assessment succeeds in attaining these objectives relating especially to the students, teachers and teaching activities, largely depends on how well-equipped the teachers are about measurement and assessment" (p. 1997). To achieve this, one needs to observe the terminology of 'perception' and 'literacy' of the teachers towards the assessment.

Language Assessment Literacy: Perception and Practice

Assessment literacy depends upon teachers' being able to use their assessment-related knowledge effectively in their contexts. As Kahl et al. (2012) describe assessment literacy as enclosing some abilities to select and build different aims such as accountability, instructional program evaluation student growth monitoring and/or promotion and diagnosis of specific student needs (learning gaps) (p. 8).

Mertler and Campbell (2005) developed an inventory, titled as Assessment Literacy Inventory (ALI), to measure teachers' assessment literacy. They established a framework that introduces the core concepts of assessment literacy including types and qualities of measures

for teachers and school administrators. They aimed at helping teachers and school districts to develop valid student grading process and benefit from that inventory in terms of *in-class assessment* and professional development. Table 1 below shows those core concepts of in-class assessment for those who are willing to become assessment literate.

Table 1

Core Concepts of Assessment Literacy for Teachers and School Administrators

Standards	Teachers must be able to create/select and effectively use classroom assessments for a variety of purposes.		Teachers and administrators must be able to select and effectively interpret and use results from external interim and summative assessments designed for a variety of purposes.
Category of Measures	Formative	Classroom Summative	External Interim and Summative
Types of Measures	• Formative assessment evidence gathering techniques	• Selected-response • Constructed-response • Performance tasks • Portfolios	• District benchmark • Diagnostic • General achievement • Adaptive • State accountability

Quality of Measures	• Unpacking standards
	• Depth of knowledge
	• Quality of evidence regarding learning targets
		• Good and bad items/tasks
		• Reliability and validity
		  · Test length
		  · Domain representation (See "Alignment")
			• Match to purpose
			• Universal Design
			• Item quality in banks and tests
			• Item selection criteria
			• Alignment
			  · Categorical concurrence
			  · Depth of knowledge
			  · Range of knowledge
			  · Balance of representation
			• Technical characteristics (reliability, validity)

Results and Their Use	• Quality and use of feedback
	• Use of data to inform instruction
		• Scores vs. grades
		• Effective and detrimental grading practices
			• Reporting statistics
			  · Scaled scores
			  · Percentile ranks
			  · Performance levels
			• Subgroup/subtest results
			• "Growth" and longitudinal data
			• Comparability issues

Apparently, understanding those core concepts of assessment literacy may benefit for teachers especially in order to have better, valid and reliable outcomes, for that reason, Popham (2006) suggests, "…educators—and everyone else who has an interest in education—need a dose of *assessment literacy*" (p. 84). When analyzing assessment literacy, one needs to have a look at pre-service programs' efficiency on teacher assessment literacy since teacher candidates start learning the terms of 'assessment' or 'in-class assessment' at the Bachelor years. Concordantly, the studies show that pre-service program education is not enough for teaching assessment to candidate teachers. Malone (2008) claims that "…there is no consensus on what is required or even needed for language instructors to reliably and validly develop, select, administer and interpret tests" (p. 225). Shepard & Sheppard (2000) support the inefficiency of pre-service program and say, "In traditional educational measurement courses, pre-service teachers learned about domain specifications, item formats,

and methods for estimating reliability and validity. Few connections were made in subject matter methods courses to suggest ways that testing might be used instructionally" (p. 1).

Gok et al. (2012) conducted a research on 200 pre-service elementary and preschool teachers about their perceptions when it comes to assessment by observing metaphor analysis; they were asked to complete the sentence such as "measurement and assessments is like… because…". The results showed that half of the teachers chose the positive perceptions and the other half chose the negative associations towards assessment. But in most common sense, their perception towards assessment is negative. Also big number of teachers perceives assessment as 'numerical data'. In that sense, assessment is associated with negative adjectives and 'complementary side' of it is neglected since the participants perceive it as 'numerical data'. As a result, in assessment process, these misperceptions may affect students' performances negatively.

Brookhart (2001) states in her article that teachers are not competent enough to use valid grading practices and interpret standardized tests. At the same time, Stiggings (2001) states that there are numbers of teachers and administrators who have low level of assessment literacy and that level of literacy cause students to be assessed inaccurately and avoid them to show their full potential (as cited in Mertler & Campbell, 2005, p. 4).

Muñoz, Palacio, & Escobar (2012) conducted a study on 62 Colombian teachers beliefs about assessment in general and whether those beliefs match with their practices or not. They used surveys, written reports and interviews as instruments. The results demonstrate that there is a mismatch between what they believe and what they practice which suggests some more opportunities for teachers to be able to practice their assessment skills in educational settings.

Rueda and Garcia (1994) examined 3 groups of third and fourth grade teachers' (special education pull-out, bilingual credentialed and bilingual waivered) belief systems and everyday practices in terms of reading assessment on Latino language minority students in California. The study was conducted by the help of semi-structured interview, a written questionnaire, and classroom observation. Results show that there is a great deal of variance in 3 different groups of teachers' assessment belief system and practices. Also, there is a general discrepancy between the beliefs of the teachers and many of the new educational initiatives in assessment and instruction.

Mertler (2003) conducted a study in order to learn about pre-service and in-service teachers' 'assessment literacy' and 67 pre-service teacher candidates who just completed their course on classroom assessment at university participated in the study. The other half of the participants was in-service teacher including 197 teachers who were working in different district schools. The researcher used Classroom Assessment Literacy Inventory or CALI consisting 35 items. Results showed that traditional teacher preparation courses are not realistic and they are not in concordance with preparing teachers to real classroom practice. Also he underlines the output of pre-service education on teacher assessment, as "They often believe that they have not received sufficient training in their undergraduate preparation programs in order to feel comfortable with their skills in making assessment decisions" (p. 22).

Jannati (2015) conducted a study on 14 EFL instructors according to their experiences as low, mid and high experienced teachers. The study was conducted via semi-structured interview. The aim of the study was to investigate the extent to which EFL instructors reflect their assessment literacy to their practices and whether experience effect assessment literacy and practice. The results showed that experience do not have effect on their assessment literacy. Also although the participants were assessment literate they were not able to reflect their literacy to their practices.

Considering Turkish context in terms of assessment in classes, studies related to this area are limited (Hatipoğlu 2017). Even so, it is evident that in Turkey, the system related to testing has changed a lot recently (6 times since 2003) and it is crucial for language teachers to assist well to students to be successful in school context. In relation to this, Hatipoğlu (2015) conducted a study on 124 pre-service teachers between the years of 2009 and 2012 at Middle East Technical University (METU) in Turkey with the aim of understanding what those future ELT teachers know about language testing and also what their expectations are in their departmental course of English Language Testing and Evaluation (ELTE) in terms of content and methodology. The study showed that, almost all the future ELT teachers expected to evaluate, select and write exams and finally prepare students for the national and international examinations with the knowledge coming from ELTE course. Also, local assessment context in Turkey (being exam focused) shape pre-service teachers beliefs in ELTE department course at university (Hatipoğlu 2016). Interestingly, after four years of training in ELT in METU, students still had very limited knowledge related to testing.

Hatipoğlu (2010) worked with 81 ELT department students in order to obtain a summative evaluation of an undergraduate English Language Testing and Evaluation (ELTE) course delivered at the Department of Foreign Language Education in the Faculty of Education at METU in Ankara and also whether this course help those undergraduate students to prepare them to face the challenges in terms of assessment in their future careers. Questionnaires and interviews were used to collect data. The results revealed that, the participants selected 3 topics that they thought would help them in their future careers, which were testing skills/knowledge, reliability and validity. Another issue was, based on the students' feedback, additions were required in ELTE course. Students expressed their need for the creation of more time and opportunities for practicing writing and evaluating various exams. Another valuable finding was students found some of the sessions of ELTE course too abstract and it was hard for them to make sense. In this way, they asked for more presentations prepared by them and presented to their classmates that would make them more involved and understand the concepts better.

Assessment tools have big role in shaping language teachers' instructions in classes as well. In 2012, Haznedar conducted a study on 538 primary school teachers in order to understand what kind of assessment tools was used in order to evaluate those kids' success. According to the results of her study, it has been identified that, nearly all of the participants (98.9%) use out-of-date written examinations and multiple-choice questions (78.2%) in order to evaluate primary school children's performance. According to Köksal (2004), there is insufficient training in terms of assessment meanly in testing and when the word 'test' is uttered, 'multiple choice test' comes immediately to the mind of most of the language

teachers in Turkey. According to Frost (2005) underlines the negative side of using one assessment tool, meanly tests as, "they are 'one-off' events that do not necessarily give an entirely fair account of a student's proficiency".

Overall, considering the theoretical part of the study, it can be inferred that, there is inefficiency about assessment by the teachers due to some reasons; (1) pre-service education is insufficient in assessment courses, (2) there is a mismatch between perception and classroom practice and finally (3) assessment literacy does not reflect real classroom practice.

In conclusion, it is important to search about teachers' beliefs regarding their in-class language assessment and how they reflect their beliefs into practice. Studies regarding the concept of assessment literacy; there is not enough data for making more inferences about assessment literacy and its practice in Turkey. So this study was conducted to examine Turkish EFL instructors' beliefs on in-class language assessment and reflection of their belief system on their classroom practice.

Methodology

Research Questions

The current study was conducted with the aim of understanding EFL instructors' perceptions towards *in-class language assessment* (while they are teaching) and its reflection in their classroom practice for the sake of learning. The study also investigated the relationship between experience and perception of in-class language assessment. To achieve these, following questions were proposed:

1. What are Turkish EFL instructors' perceptions of in-class language assessment?

2. How do Turkish EFL instructors reflect their assessment literacy in their own practice?

3. If any, is there any relationship between experience and perception of in-class language assessment?

Participants

In this study, data was collected from twelve instructors, 8 (female) and 4 (male) preparatory school instructors working in English Preparatory Program of a Turkish university. They have different educational backgrounds such as English Language Teaching, English Language and Literature, and English Translation and years of experience in teaching. Also, the participants' experiences range from 1-15 years. Taking participants' experiences into account, the participants were classified as novice teachers (NT) and expert teachers (ET). "The term 'novice teacher' is defined as a teacher with less than 5 years of teaching experience" (Kim & Roth, 2011, p. 4). In this sense, regarding years of expertise, 7 of them were NTs and 5 of the instructors were ETs in the related educational setting.

Data Collection Tools and Analysis

A qualitative research method was implemented and Phenomenological research design was used for data collection and analysis in order to seek to discover the perception and practice of instructors regarding their in-class assessment framework. Phenomenological research design was chosen for this study since it aims to examine instructors' lived experiences as they perceive those experiences provided by them. For the study, semi-structured interview was applied in order to collect data (see Appendix A). The interview was developed by Jannati (2015) and was also conferred to the academic advisor. The interview has 17 open-ended questions. The aim of the semi-structure interview was to collect detailed data about how Turkish EFL instructors perceive in-class language assessment and how they demonstrate that knowledge to their classroom practices. Since the university was giving standardized tests as an assessment tool, the focus of this study was on teachers' perceptions on *in-class assessment* and their practices taking their perceptions into account. When mentioning about in-class assessment, presentations, portfolios, group discussion, projects, individual conferences, cooperative tasks, oral recordings, think-aloud and even attendance to the lecture are meant to be considered meanly formative assessment tools.

At the beginning of the research, very general information was given to the participants about in-class assessment before starting to the interview. Then mentioned interview was conducted to each participant individually. Each interview was recorded to a smart phone as an audio file to be listened later on. After the interviews, the audio files were listened, transcribed and then analyzed related to its content.

Results

In the following part of the study, teachers' perceptions and their practices are analyzed under two different titles to see whether experience plays an essential role while shaping the concept of 'in-class assessment'.

Instructors' Perceptions on Assessment

In the study, all the participants stated the importance of 'assessment' in a language program in order to get feedback from the students and to keep track on their progresses for the question of "Do we need assessment in a language program?". One of the participants, who is ET, with experience of 14 years comment was eye-catching since that instructor complained assessment to be too numerical:

• "We need an assessment in a language program because they can at least understand in which path they are, but maybe not too much numerical because numbers are dominant in assessment process I think". (ET)

Clearly, one of the ET understand the substantial part of the assessment but at the same time seems to be sorry for the focus on numerical side of the assessment.

For the following question of "What are the characteristics of good assessment?", the participants defined some characteristics for what a 'good assessment' is and the ideas and definitions varied Below the table, there are some characteristics specified for a ' what good

assessment is' by the participants and their frequency numbers are shown next to each stated characteristic.

Table 2

Characteristics of a 'good assessment' defined by the instructors

Characteristics	NTs	ETs
Fair	1	2
Valid	4	-
Effective	1	-
Assessing all the skills	-	1
Reliable	4	2
Equal	1	1
Organized	1	-
Product-based	2	-
Authentic	1	-

Table 2 shows that most common uttered characteristics for a 'good assessment' from the participants were assessment's being *valid* (it is mentioned 4 times by the instructors) and *reliable* (mentioned 6 times by the participants) which indicates assessment to assess what it is supposed to assess (valid) and assessment that brings the same results in repeated different times (reliable). As it is shown in the table, it is clear that all the participants have some basic definition for the quality of a good assessment. So all the instructors, regardless of their experiences, agree on the necessity of assessment in language programs and can basically define a 'good assessment' with their knowledge of teaching.

Asking the participants opinions on "Should students be informed of what they will be assessed on?", all the ETs indicated that students should be informed of what they will be assessed on beforehand because they think that it is their right to know. But for the NTs, there were some diverse ideas about the question; some of the ideas are as follows:

• "They should be informed most of the time so that they will feel much better and confortable, they maybe nervous the other way." (NT)

• When they know that they are going to be assessed, they will be prepared for it and I do not want that. They have to be prepared all the time because it is a language, it is not a set of formulas." (NT)

• "Not always. If they will be graded, they need to be informed that they will be assessed like mid-terms, exams and quizzes; more likely for formal assessments." (NT)

- "If I am going to give grades, yes they need to be informed but if I am not going to give grades, they do not need to be informed." (NT)

Half of the NTs thought it would be nice to inform students before so that they will know what to expect and take it seriously also they get mentally prepared for it. But for the other half of the NTs, it was partly acceptable; it was acceptable if the assessment is formal assessment like mid-term and final exam. But other than that, if it is in-class assessment, it is not necessary to inform students for every single activity.

The following question was for the assessment rubric: "Should students be informed of the assessment rubric?". Not taking experience into consideration, 10 out of 12 participants stated the importance of the informing the students about the assessment rubric. Additionally some NTs stated that at the very beginning of the year, they inform the students about the criteria like portfolio, presentation and group projects type of assessment criteria so that they will know what they will face with and prepare themselves accordingly.

When asked about "Do you think students' scores represent what they have learned?" to the instructors, as opposed to ideal expectations, all the 12 participants pointed that students' scores do not match with what they have learned in the classroom. Some of the feedbacks are as follows:

- "Not really. When students prepare for the exams, they only focus on grammar and the rest of the skills are being ignored. Later on, they come and complain about it!" (ET)

- "No. Especially for speaking, sometimes they are so nervous, even if they can express themselves very well in my classes, they can't perform well in the exam since they get so nervous." (NT)

- "Never. I cannot fully tell that score fully represents that specific student. One day I came across with a placement test; the student did 98 out of 100 and when we did the speaking exam, he was A1! So considering my experience, there is a mismatch with the scores and learning." (ET)

Actually those statements carry great importance since they reflect the overall state of the whole language education system. Apparently, there is also a mismatch between students' scores and performance in their classes which is another big issue to pay great attention.

Focusing on the question of "When are you satisfied with assessment results?", regardless of the experience again, 5 out of 12 participants get satisfied when they see that their students actually *produce* the language namely using the language that they learn in the classroom. 3 of the participants get satisfied when their students' scores increase in the exam, which shows that those instructors have numerical expectations actually it is also what the system requires as well. The rest of the participants' answers varied as:

- "I feel satisfied when I see my students correct their own mistakes" (NT with 3 years of experience)

- "Satisfaction comes alive when I see my students feel independent in using the language and also be creative about the language." (ET with 14 years of experience)

Seeing a production in English makes the teachers fulfilled which also shows that the instructors in the study see the language English as a live thing and that's why, productive skills are crucial for them.

For the question of "Does assessment impact the way students study and do assignments?", 12 out of 12 participants confirmed the idea of assessment impact on students' studies and assignments. Regardless of experience, 4 of them stated that assessment, as an idea, motivated them. The others expressed that since that is a system requirement, students take it seriously and do it. Some extracts shown below:

- "I really wish that they do it for themselves but all the time that really works is the grades! They have to be there." (NT)

- "Yes because they see that they will get grades from the homework activity or exam, they study with a great care and pay attention to most of the details which make them to be motivated as well." (ET)

Those limited comments also display the situation of the education system as well since the end of the term matter to a great deal for students; they do not have fun during the journey (learning process) but just spend their time by thinking about the result (the total score) meanly the end of the journey. That's why, English is perceived as monotonous lesson just memorizing the rules of grammar and passing from exams with flying colors.

Not last but least, when asked "To what extent are you convinced that your assessment is valid/fair?", regardless of their experiences, 3 participants stated that seeing same output both from in-class performance and exam results while 4 participants indicated that they are convinced when they see production. The rest of the 5 participants expressed that they are convinced when there is no difference among the students' scores.

In conclusion, considering the perception of instructors in terms of assessment, most of the answers did not show any experience effect except informing students earlier about the assessment. ETs agreed on informing students earlier, it maybe because they demand transparency; if everything is known earlier by the students, the instructors will not have difficulty later on. Overall, there is no experience effect in shaping teachers' perceptions on assessment.

Instructors' Practices on Assessment

The timing of the assessment in the classroom is vital. For the question of "When do you assess your students?", the results show that there are basically two types of assessment

come alive; summative and formative assessment. All the ETs conveyed that they use formative assessment since learning is a process even outside the classroom that's why; there is not specific time for assessment. For the NTs, 5 out of 7 asserted they use formative assessment. Only 2 of the NTs stated that they use summative assessment such as at the end of the activity or lecturing. In Table 3, there are some in-class tools used by the participants in order to assess their students during the classroom time and also their frequency numbers are given below.

Table 3

Number of the In-Class Assessment Tools used by the Instructors

Tools	NTs	ETs
Pair-work activities	3	2
Teacher's book instructions	1	-
Speaking-related activities	2	4
Worksheets	2	1
Presentation	1	1
Portfolio	1	-
Group work	2	1

The most common used tool is speaking related activities since they stated that, in this way they would be able to use the language. The other tool mostly used is pair-work activity since 5 of them stated that, that kind of activity makes them feel comfortable and gives a feeling of equality. The data shows that, tools used by all the teachers are basically tools that would make students to interact with each other with the help of pair work and speaking focused activities.

"Which skills and language components do you typically assess?" was a critical question that would give a lot clues about the focus of the teachers. 10 out of 12 participants responded that they put emphasis more on speaking as a productive skill to be assessed in the classroom as considering the language a live thing, which should be used in the classrooms as often as possible. 1 NT indicated that his emphasis is more on pragmatics in the class so students will be able to use the language appropriately. The other NT stated that the emphasis is more on grammar since it is a system requirement. These indicate that most of the teachers in the study are aware of the importance of the 'speaking' skill and enhance their students accordingly but the focus for assessing students is mostly on speaking which is not enough in order to see a student as a whole since there are 3 more skills to be considered.

Peer and self-assessment are two current concepts in classroom assessment (Khonbi & Sadeghi, 2013, p. 1553; Boud, Cohen, & Sampson, 2014, p. 1). Regardless of their

experiences, 2 out of 12 teachers stated that they use peer and self-assessment very rarely since it would not be nice to make them assess or give feedback about each other, in those teachers' perspective, students do not feel comfortable while commenting on the partner's work. 8 teachers prefer peer assessment since it would be easier to detect their friends' mistakes easily and they would feel more comfortable with their pairs. On the other hand, when the teacher assesses the students, the students do not feel very comfortable about being assessed by an authority and it may affect students' performance negatively.

- "I prefer peer-assessment in my class most of the time because they feel that they are building trust to each other and it makes them be critical to each other as well." (ET)

- "They do pair works and group works; I ask them to talk about their friends. I observe that they feel much more confident as a result." (NT)

Also they stated that peer-assessment helps the teacher to use time effectively since there is a lot of workload for the teachers and size of the classes may be crowded to give feedback for everybody at the same time. The other 2 teachers prefer to use self-assessment so that it will not cause a problem in the class such as listening to students complaining about giving lower grades on purpose etc.

Not taking experience into consideration, 9 of 12 teachers claimed that they do not consult their students in shaping their *assessment* for the question of *"To what extent do you consult your students about assessment?"*. One of the ET (14 years of experience) indicated as:

- "They are the people who are assessed so it would be much more fair. But it is difficult to do it since administration part limits you too to shape your assessment according to students' feedback." (ET)

The rest of the participants (3) consulted their students about his/her in-class assessment at the very beginning of the year.

Using assessment results for the sake of students helps students overcome the problematic parts of the language and makes learning better (Campbell & Levin, 2009, p. 48). All the 12 instructors pointed out that after collecting some data (such as asking questions about the related topics to the students, in-class tasks like presentations, games or mini revision tests) after in-class assessment, if there is a problematic part or a part that needs to go back and have a look at again, they prepare a remedial work in order to cover that specific part before moving on. 6 teachers expressed that they make some revisions and prepare games for mistaken parts. Other 4 participants stated that they give more homework in order to repair the ill parts and finally 2 participants mentioned changing some techniques accordingly.

Using a table of specification (TOS) to plan assessment helps teacher in the in-class assessment process. As Notar et al. (2004) mentioned, "A Table of Specifications identifies not only the content areas covered in class, it identifies the performance objectives at each level of the cognitive domain of Bloom's Taxonomy." (p. 117). For the question of 13, "Do you use a table of specifications to plan assessment?", most of the NTs and ETs (10 teachers) do not use a set of table for their classrooms while assessing their students during the classroom period, which means that they depend on their own judgments while assessing their students. One of the NTs stated that she makes use of CFR descriptors in order evaluate her students and the other teacher mentioned that he uses ready-made criteria for his students so it shows that the participant at least has a basis, something valid, to assess the class during teaching.

All the teachers should have some knowledge on assessment since it is an important part of the learning and teaching process as stated by all the participants. That way, they will have more organized and objective judgments on their decisions. Nearly all of the participants do not act specifically to learn something about classroom assessment apart from searching online communities and engaging in informal office talks.

Discussion and Conclusion

The study aimed to find out instructors' literacy on in-class assessment and its' reflection in their practices. Since teachers have an important role in decision-making in assessment, there is a growing demand on teacher as well. The findings, in general, revealed that there is a mismatch between Turkish EFL instructors' in-class assessment literacy and its reflection in the practice. The results above were tried to be explained more broadly in this part of the study.

For the instructors' perception part of the study, the results show that instructors, regardless of their experiences, are aware of the importance of the in-class assessment let alone assessment in general, which shows that they all agree on the necessity of it. At the same time, they are also familiar with some very basic components such as definition and characteristics of in-class assessment but only NTs were able to use variety of terminologies. It is probably because most of them (5 NTs) were studying English language teaching in MA program and also they were newly graduate from Bachelor's Degree. This shows that being able to use variety of terminologies and having theoretical knowledge about in-class assessment is highly related to being recently graduated but other than that with the experience, the terminology and theoretical knowledge transform into practical knowledge for in-class assessment.

One of the distinctive finding of the study was that all the instructors agree that students' scores do not match with what they have learned in the class. It may be due to the Turkish education system. According to Boyacıoğlu (2014), "A majority of Turkish students are unable to improve their English language skills due to poor language training methods at state schools, according to a joint report from the British Council and the Economic Policy Research Foundation of Turkey (TEPAV)" (para.1). As it is stated, one of the reasons may be because of the poor techniques. This noteworthy data shows that teachers should re-consider

their techniques and testing units would better take classroom learning into consideration when preparing exams.

From the instructors' practice part, there are some important inferences to be made. It was clear that even though instructors are, to some extent, assessment-literate however it was hard for them to reflect their knowledge in their practices. For instance, the instructors define some characteristics for a good assessment such as *valid*, *reliable*, *authentic* and *assessing all the skills but* when asked about their practice, there were very different answers about their own validity and reliability, at the same time, most of them only focused on speaking skill when evaluating their performance. This demonstrates a mismatch between perception and practice. Also the in-class assessment tools used by the instructors in the class are limited. On the contrary, there are variety of in-class assessment tools which can be used such as group discussion, projects, individual conferences, cooperative tasks, oral recordings, think-aloud, games, question-answer, story telling, songs tasks and etc.

Another substantial finding is that all of the ETs and half of the NTs use formative assessment. Garrison & Ehringhaus (2007) illustrate the importance of formative assessment as, "…it provides the information needed to adjust teaching and learning while they are happening. In this sense, formative assessment informs both teachers and students about student understanding at a point when timely adjustments can be made" (p. 1). It can be inferred that formative assessment is mostly preferred type of assessment in the class for the teachers which does not only take tests and exams into consideration but takes the whole learning inside the classroom into consideration as well.

Nearly all the participants do not use any table of specification (TOS) while evaluating the students' in-class performance. This shows that teachers heavily depend on their own instinctive judgments while observing and assessing their students' performance. That's not based upon a scientific theory since judgments depend on person to person. This is not a trustworthy attitude towards the students' performance from the perspective of both students and school principals. In this case, teachers are in need of a TOS suggested by the administration in order to increase the reliability of their in-class assessment.

Overall, there is a mismatch between perception and practice of in-class assessment but there is not experience effect observed when it comes to the mismatch between perception and practice. It can be said that this study has pedagogical implications for institutions to provide some continuing professional development sessions to the instructors to raise their awareness and to improve their practices about their own assessment, which covers big part of teachers' teaching process. As it is stated by Stiggins (1988), "Classroom assessment requires a great deal of time and effort; teachers may spend 40 % of their time directly involved in assessment related activities" (p. 363). Also Kahl et al. (2012) suggest "…we expect that any effective pre or in-service program would require at least one assessment course involving instructors with both measurement expertise and practical experience" (p. 4).

In pre-service programs, it is vital for teacher educators to give some more space for 'assessment' as a whole in order to improve the quality of teacher assessment and better results about their practice process. As Kahl et al. (2014) suggest,

Promote candidates' mastery of assessment literacy knowledge and competencies in pre-service programs by including separate course work focused on assessment, embedding assessment topics in content and methods courses, and providing opportunities to enable candidates to apply what they have learned (p. 3).

In conclusion, the current study was conducted with the aim of understanding Turkish EFL instructors' perceptions towards language assessment and its reflection in their classroom practice. The findings of the study show that, although most of the teachers were familiar with basic classroom assessment, when it comes to classroom practice, there is an imbalance between assessment literacy and classroom reflection. It also revealed that there is not much relationship between the experience and assessment perception.

Further research can be conducted in a different educational context, meanly in high school or secondary school in order to find much more different, eye-catching data since the current study was conducted at a private university's preparatory program. Also the study was conducted on only 12 EFL instructors If it is conducted with a larger teacher population, the data will vary. Another limitation of the study is that it was only conducted via semi-structured interview. In order to validate those data, it may be beneficial to use different data collection tools. In that way, the applicability and influence of those findings may be wider and stronger.

References

Boud, D., Cohen, R. & Sampson, J. (1999) Peer Learning and Assessment, Assessment & Evaluation in Higher Education, 24:4, 413-426.

Boyacıoğlu, H. (2014, March 17). Turkish students fail to go beyond 'hello' in English courses. *Hürriyet Daily News*. Retrieved from http://www.hurriyetdailynews.com/turkish-students-fail-to-go-beyond-hello-in-englishcourses.aspx?pageID=238&nID=63660& NewsCatID= 341

Brookhart, S. M. (2001). The Standards and classroom assessment research. Paper presented at the annual meeting of the American Association of Colleges for Teacher Education, Dallas, TX. (ERIC Document Reproduction Service No. ED451189)

Campbell, C., & Levin, B. (2009). Using data to support educational improvement. Educational Assessment, Evaluation & Accountability, 21(1), 47-65. http://dx.doi.org/10.1007/s11092-008-9063-x

Coombe, C., Folse, K., & Hubley, N. (2007). *A practical guide to assessing English language learners*. Ann Arbor: MI: University of Michigan Press.

Fleming, M. (2007). *Languages of schooling within a European framework for language of education: Learning, teaching, assessment. – Report.* Report of the Intergovernmental conference of the same name, Prague, 8-10 November 2007. Strasbourg, France: Council of Europe, Language Policy Division.

Frost,   R.   (2005).   *Testing   and   Assessment.*   Retrieved   from https://www.teachingenglish.org.uk/article/testing-assessment

Garrison, C., & Ehringhaus, M. (2007). Formative and summative assessments in the classroom. Retrieved from

    F http://www.amle.org/Publications/WebExclusive/Assessment/tabid /1120/Default.aspx

Gök, B., Erdogan, O., Altinkaynak, S., & Erdogan, T. (2012). Investigation of Pre-Service Teachers' Perceptions about Concept of Measurement and Assessment Through Metaphor Analysis. *Procedia-Social and Behavioral Sciences*, *46*, 1997-2003.

Hatipoğlu, Ç. (2010). Summative Evaluation of an Undergraduate English Language Testing and Evaluation Course By Future English Language Teachers. English Language Teacher Education and Development (ELTED), 13 (Winter 2010), 40-51.

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.

Hatipoğlu, Ç. (2016). The impact of the University Entrance Exam on EFL education in Turkey: Pre-service English language teachers' perspective. *Procedia-Social and Behavioral Sciences,* 232, 136-144.

Hatipoğlu, Ç. (2017). History of English language teacher training and English language testing and evaluation (ELTE) education in Turkey. In Yasemin Bayyurt and Nicos S. Sifakis (Eds). *English Language Education Policies and Practices in the Mediterranean Countries and Beyond* (pp. 227-257). Frankfurt: Peter Lang.

Haznedar, B. (2012). Perspectives on pre-service education lf English language teachers in Turkish primary schools. In Y. Bayyurt & Y. Bektaş-Çetinkaya (Eds.), Perspectives on Teaching and Learning English in Turkey (pp. 39-57). Berlin & New York: Peter Lang.

Jannati, S. (2015). ELT Teachers' Language Assessment Literacy: Perceptions and Practices. *The International Journal of Research in Teacher Education*, *6*(2), 26-37.

Kahl, S. R., Hofman, P., & Bryant, S. (2012). *Assessment literacy standards and performance measures for teacher candidates and practicing teachers.* Measured Progress.

Khonbi, Z. A., & Sadeghi, K. (2013). The effect of assessment type (self vs. peer) on Iranian university EFL students' course achievement. *Procedia-Social and Behavioral Sciences*, *70*, 1552-1564.

Kim, K. A., & Roth, G. L. (2011). Novice teachers and their acquisition of work-related information. *Current Issues in Education*, *14*(1).

Kondal, S. (2015). The course content and materials used in Oral Communication Skills II course: students' perspectives. *ELT Research Journal*, *4*(4), 240-254.

Köksal, D. (2004). Assessing teachers" testing skills in ELT and enhancing their professional development through distance learning on the net. Turkish Online Journal of Distance Education (TOJDE), 5(1), 1-11(https://tojde.anadolu.edu.tr/tojde13/pdf/koksal_pdf.pdf).

Malone, M. E. (2008). Training in language assessment. In *Encyclopedia of language and education* (pp. 2362-2376). Springer US.

Mertler, C. A., & Campbell, C. (2005). Measuring teachers' knowledge & application of classroom assessment concepts: Development of the" Assessment Literacy Inventory". Online Submission.

Mertler, C. A. (2003). *Preservice Versus Inservice Teachers' Assessment Literacy: Does Classroom Experience Make a Difference?*. Paper presented at the annual meeting of the Mid-Western Educational Research Association, Columbus, Ohio.

Montee, M., Bach, A., Donovan, A., & Thompson, L. (2013). LCTL Teachers' Assessment Knowledge and Practices: An Exploratory Study. *Journal of the National Council of Less Commonly Taught Languages*, 13,1-31.

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

Notar, C. E., Zuelke, D. C., Wilson, J. D., & Yunker, B. D. (2004). The table of specification: insuring accountability in teacher made tests. *Journal of Instructional Psychology*, *31*(2), 115.

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

Popham, J. W. (2006). Needed: A Dose of Assessment Literacy. *Educational Leadership*, *63*(6), 84-85.

Rea-Dickins, P. (2004). Understanding teachers as agents of assessment. *Language Testing,* 21, 249-258.

Rueda, R., & Garcia, E. (1994). Teachers' beliefs about reading assessment with Latino language minority students. Center for Research on Education, Diversity & Excellence. NCRCDSLL Research Reports. Berkeley, CA: University of California.

Rust, C. (2002). Purposes and principles of assessment. *Learning and Teaching Briefing Papers Series*. Oxford: Oxford Brookes University, Oxford Centre for Staff and Learning Development OCSLD.

Salvia, J., Ysseldyke, J., & Witmer, S. (2012). *Assessment: In special and inclusive education*. Cengage Learning.

Shepard, L. A., & Sheppard, L. A. (2000). *The role of classroom assessment in teaching and learning.* . Santa Barbara, CA: Center for Research on Education, Diversity & Excellence, University of California, Santa Barbara.

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

Stiggins, R. J. (2001). The unfulfilled promise of classroom assessment. Educational Measurement: Issues and Practice, 20(3), 5-15.

Susuwele-Banda, W. J. (2005). Classroom assessment in Malawi: Teachers' perceptions and practices in mathematics. Unpublished doctoral dissertation. Virginia Polytechnic Institute and State University, Virginia.

Thomas, J., Allman, C., & Beech, M. (2004). *Assessment for the diverse classroom: A handbook for teachers.* Tallahassee, FL: Florida Department of Education, Bureau of Exceptional Education and Student Services. Retrieved from http://www.fldoe.org/ese/pdf/assess_diverse.pdf

Yamtim, V., & Wongwanich, S. (2014). A study of classroom assessment literacy of primary school teachers. *Procedia-Social and Behavioral Sciences*,*116*, 2998-3004.

Appendix A

Semi-structured interview questions

1) Do we need assessment in a language program?

2) What are the characteristics of good assessment?

3) When do you assess your students?

4) How do you assess your students?

5) Which skills and language components do you typically assess?

6) Do you engage students in peer- and self-assessment? If yes, how?

7) To what extent do you consult your students about assessment?

8) Should students be informed of what they will be assessed on?

9) Should students be informed of the assessment rubric?

10) Do you think students' scores represent what they have learned?

11) When are you satisfied with assessment results?

12) How do you make use of assessment results?

13) Do you use a table of specifications to plan assessment?

14) Does assessment impact the way students study and do assignments?

15) To what extent are you convinced that your assessment is valid/fair?

16) Do you support the idea that instructors need to have some sort of background knowledge about   classroom assessment?

17) How do you increase your knowledge about assessment? Do you read books, attend

workshops…