Private School EFL Teachers' Perceptions On Assessment: A Case Study In Turkey

Okan Önalan
Independent Researcher

Esim Gürsoy
Bursa Uludağ University

ABSTRACT: Teachers' selection and use of appropriate assessment methods requires not only their contemplation on lesson objectives, but also their knowledge, awareness and understanding of assessment processes. The present study aims to investigate Turkish private school language teachers' perceptions, use and understanding of assessment as an integral part of their teaching practices. Using a sequential mixed method approach, researchers triangulated the data to get an in-depth understanding of the phenomena: a 5-point Likert type online survey, participants' self-reports on assessment terminology as a part of an in-service training session, and participants' metaphors to explain formative and summative assessments which were collected during the same training. Results revealed participants' high self-confidence on traditional summative uses of assessment, but moderately low perceptions on using assessment formatively. Findings showed significant inconsistency between participants' limited knowledge on assessment terminology as opposed to high self-perceptions.
Key words: EFL assessment techniques, formative assessment, perceptions on assessment, summative assessment, metaphors

Percepciones de profesores de inglés como segunda lengua (EFL) de escuelas privadas sobre la evaluación: un estudio de caso en Turquía.

RESUMEN: La selección y el uso de los métodos de evaluación adecuados por parte de los profesores no solo requiere que consideren los objetivos pedagógicos, sino también que conozcan, comprendan y sean conscientes de los procesos de evaluación. El presente estudio tiene por objetivo investigar las percepciones de los profesores de EFL de escuelas privadas turcas y su uso y la comprensión de la evaluación como parte integral de sus prácticas pedagógicas. A través de métodos mixtos secuenciales, los investigadores triangularon los datos para entender a fondo los fenómenos: una encuesta por Internet basada en la escala de Likert con 5 puntos, autoinformes de los participantes sobre la terminología de la evaluación como parte de una sesión de formación de profesores en ejercicio activo, y metáforas de los participantes para explicar evaluaciones formativas y sumativas recopiladas durante la misma formación. Los resultados revelaron una alta autoconfianza de los participantes en el uso sumativo tradicional de la evaluación, pero percepciones moderadamente bajas sobre el uso formativo de la evaluación. Los hallazgos mostraron una inconsistencia significativa en el conocimiento limitado de la terminología de la evaluación entre los participantes, a diferencia de las altas autopercepciones.
Palabras clave: técnicas de evaluación de EFL, evaluación formativa, percepciones sobre evaluación, evaluaciones sumativas, metáforas.

1. INTRODUCTION

Assessment plays a fundamental role in instructional processes. Literature is abundant with assessment terminology, e.g. formative, summative, large-scale, low-stakes, alternative assessments etc., which sometimes might overwhelm teachers rather than giving them sufficient guidelines on how to effectively use assessment as an instrument to enrich their teaching. In such cases, teachers tend to become reluctant to make use of assessment as a tool for better instruction and they stick to traditional uses restricted to exams and grades. This limited approach rules out elements such as differentiated feedback, diagnostic needs analysis and progress testing which are otherwise beneficial means of improvement in student understanding. Therefore, teacher training programs should go beyond delivering mere definitions of assessment terms and focus more on including modules or courses that educate teachers on how to use the assessment factor effectively in their instruction.

As in most cases in education, the purpose of teacher training programs should not only be achieving better understanding of a given subject and its hypothetical use in the classrooms, but also developing teachers' ability to transfer that perceived understanding to authentic and novel situations for improved instruction. However, especially pre-service teacher training programs in Turkey lack the effective implementation of such a component (Balbay, Pamuk, Temir, & Doğan, 2018). Such inefficiency exists in terms of assessment, resulting in most novice teachers' graduating from their departments without adequate knowledge and skills to use assessment efficiently in their teaching (Ölmezler-Öztürk & Aydın, 2019). What is more, in-service training (INSET) programs for teachers generally fall short in offering the necessary training that will mitigate the problem (Balbay, Pamuk, Temir, & Doğan, 2018; Demirtaş, 2008). In order to better address the issue, teachers' perceptions and implementation of educational assessment need to be closely investigated. Many research studies on the issue mainly focus on describing teachers' perceptions or understanding of assessment. Therefore, there is a need for delving further into the teachers' cognition via a qualitative perspective during which the teachers can display their comprehension. With regards to this gap in the literature, the present study aims at investigating privately-owned elementary school English teachers' perceptions, use and understanding of assessment and assessment terminology in a sequential mixed-methods research design; thus, the study is significant as it touches an under-researched issue via data triangulation.

2. LITERATURE REVIEW

Since assessment can be used in different ways, e.g. summatively (assessment of learning) and formatively (assessment for learning) (Earl, 2003), depending on the instructional setting and lesson objectives, teachers should be aware of and knowledgeable about the various functions and purposes it can serve. That way teachers can take advantage of various benefits assessment brings to educational environments (Balagtas, Dacanay, Dizon, & Duque, 2010; Farley-Ripple, Jennings, & Buttram, 2019). Although such benefits are obvious to educational theoreticians and researchers, literature suggests that many teachers, novice teachers in particular, still lack the knowledge, skills, expertise and self-confidence in integrating assessment practices into their teaching (Owen, 2016; Popham, 2009; 2013;

Yamtim & Wongwanich, 2014). Furthermore, there may be significant differences on how teachers use various assessment approaches (DeLuca, Valiquette, Coombs, LaPointe-McEwan, & Luhanga, 2018) as well as substantial discrepancy between how teachers perceive their own knowledge on assessment and how much assessment literacy they actually possess (DeLuca, LaPointe-McEwan & Luhanga, 2016; Gotch & French, 2014).

A similar inconsistency exists between how much language teachers think they transfer their assessment skills into actual assessment practices and how much they actually do. Various recent studies conducted in different parts of the world regarding teachers' practices on English language assessment have shown that language teachers may frequently fail to transfer their theoretical assessment knowledge into actual classroom practice (Crusan, Plakans & Gebril, 2016; Giraldo, 2018; Kvasova & Kavytska, 2014; Tsagari & Vogt, 2017; Vogt & Tsagari, 2014; Xu & Brown, 2016; 2017). Also, similar results have come from studies conducted in the Turkish context, where most researchers not only found relatively low level of assessment literacy on the part of teachers, but also concluded that teachers had difficulties in transferring their knowledge into practice (Hatipoğlu, 2015; Mede & Atay, 2017; Ölmezler-Öztürk & Aydın, 2019; Öz & Atay, 2017). Therefore, teacher training programs need to aim for better understanding of assessment both in the knowledge level and as a part of practical domain for teachers of all content areas. Here, a parallel concern emerges both for teacher trainers and researchers: finding adequate proof of understanding.

Understanding can be assessed using a variety of indicators. These indicators might vary from basic knowledge level skills such as giving definitions to higher-order cognitive skills such as analysis and evaluation (Anderson & Krathwohl, 2001; Bloom, 1956). The initial and the primary step of conducting appropriate assessment is prone to teachers' assessment literacy. Therefore, any investigation on teachers' perceptions of assessment should start with an inquiry of teachers' knowledge and understanding of related terminology (Malone, 2013; Vogt & Tsagari, 2014). However, this level of inquiry is never enough; nor is it a sign of teachers' understanding or ability to use the appropriate assessment methods. Thus, there is a need to look into the matter by scrutinizing teachers' deep understanding by enabling them to use higher order cognitive skills. Respectively, the ability to use metaphors is among the higher order cognitive skills that indicate understanding (Lakoff & Johnson, 1980). Therefore, 'explaining by metaphors' may provide substantial evidence about teachers' understanding and conceptualizations of assessment beyond knowledge and awareness (Zhao, Coombs, & Zhou, 2010) since metaphors can be used as tools for exploring abstract conceptual ideas and learning (Low, 2008; Yob, 2003).

Besides being beneficial means of expressing perceptions, metaphors are effective cognitive tools used constantly and involuntarily in the process of analysis (Treagust & Duit, 2015). Since metaphors present an understanding of experience, they provide holistic clues about perception (Izadinia, 2016; Wormeli, 2009). Besides, metaphors lend themselves as powerful data collection tools in qualitative research (Miles, Huberman, & Saldana, 2014; Paranosic & Riveros, 2017; Schmitt, 2005). With these factors in mind, a number of studies have used metaphors for a deeper representation of teachers' instructional philosophy (see Arslan & Karatas, 2015; Craig, 2018; Erden, 2016; Leavy et al., 2007; Saban, 2010; Saban et al., 2007). Considering that there are various factors effecting how teachers shape their instructional understanding, skills and style through time, it can be quite challenging for researchers to account for teachers' conceptualizations of teaching in general (Mellado,

Bermejo, & Mellado, 2012) as well as their perceptions of a specific domain (such as assessment) in particular. Consequently, metaphors can help teachers disclose their perceptions on classroom practices more profoundly (Erickson, & Pinnegar, 2017; Boyd & Bloxham, 2014). Therefore, the present study has used metaphors to triangulate the data and to better identify language teachers' perceptions on assessment at a higher cognitive level.

2.1. Private school EFL teaching

The global economy created a downward (from the governmental level) as well as an upward (parental demands) need for learning a foreign language (Garton, 2014; Gürsoy, Çelik-Korkmaz & Damar, 2017; Nguyen, 2011). This need has reflected itself primarily as an inclination towards English as a foreign language (EFL) in the Turkish context (Erarslan, 2019; Sarıçoban, 2012; Selvi, 2014). Unfortunately, Turkish public schools with fewer foreign language (FL) hours, fail at giving sufficient foreign language education (FLE). Hence, families who think English is a key factor in future success send their kids to privately-owned elementary schools. In line with this demand, such schools have become quite prevalent in Turkey for families who can afford high annual fees. Consequently, high expectations coming from the parents and school administrations regarding the outcome of FLE created a pressure on language teachers who work at privately-owned school chains. Unsurprisingly, it has become essential for teachers to have a good command of how, when and which type of assessment to use (Herrera & Macías, 2015; Hopfenbeck, 2019) to meet the needs of the stakeholders as well as to check learning goals. Thus, the present study was conducted to investigate private school EFL teachers' perceptions and use of assessment by focusing on how teachers that work at relatively well-funded private institutions in Turkey understand and make use of assessment as a part of their instruction.

3. METHOD

The current research has been designed as a concurrent mixed-method (QUAN + qual) study. Both quantitative and qualitative data were collected via triangulation. Initially, quantitative data were collected via "Language Teachers' Understanding of Assessment" scale (Önalan & Gürsoy, in press). Qualitative data, on the other hand, were gathered through the participants' responses on assessment terminology and their metaphors to define formative and summative assessment. While the quantitative data targeted a wider perspective on perceptions and use of assessment and assessment results, qualitative data focused on a more specific content regarding the knowledge and understanding of terms. Thus, rather than supporting the quantitative data directly as in a sequential mixed-methods research (MMR) design, a delicate and in-depth relation between understanding of basic assessment terminology and actual use of assessment was sought for by using a concurrent MMR. The current research aims to find out answers to the following research questions:

1. What are the participants' perceptions on (quantitative data);
   – involving students in the assessment process?
   – their knowledge on assessment of language skills?

– the use of assessment and assessment results?
– the selection of assessment method?
2. Are there any differences between participants' perceptions in terms of their gender and teaching experiences? (quantitative data).
3. How do the participants self-evaluate their knowledge and understanding of assessment terminology? (qualitative data).
4. To what extent the participants have the knowledge of assessment terminology? (qualitative data).
5. How do the metaphors indicate participants' understanding of different assessment types? (qualitative data).

A total of 249 ELT secondary school teachers (Female n=240, Male n= 9) working at different campuses of a private elementary school participated in the quantitative part of the study. 122 teachers had 0-5 years of experience, 93 teachers 6-11 years of experience, and finally 34 teachers 12+ years of experience. Via homogenous sampling, 20 team leaders were asked to participate in the qualitative part of the study. In homogenous sampling, participants from the same group having similar experiences were selected (Dörnyei, 2007). 20 team leaders of different campuses were invited for a face-to-face INSET on assessment and an interview. For the research purposes, this group of teachers had the knowledge of institutional goals and practices, possessed considerable knowledge about their team's (teachers') capabilities and shared similar experiences with each other. Moreover, as team leaders, one of their responsibilities was to transfer and share the content of the training with their colleagues in their own campuses. Before their participation, their signed consent was taken using an informed consent form.

3.1. Quantitative data collection

First group of data was collected via an 18-item, Likert type scale that aimed to determine ELT teachers' self-perceptions and their use of assessment prepared by Önalan and Gürsoy (in press). The instrument was validated via content, face, and construct validity analyses by the researchers. For content and face validity, the instrument was given to five experts in the field and asked to evaluate each item as 1 - not related to the construct, 2 - related to the construct but not necessary and 3 - related to the construct. As a result, content validity ratios lower than 0.99 (Veneziano & Hooper, 1997) were discarded. The same experts evaluated the face validity and some items were re-written and one item was added (see Önalan & Gürsoy, in press).

For construct validity, explanatory factor analysis was used. Principle component analysis and direct oblimin rotation were applied. Kaiser-Meyer Olkin (KMO) value (0.87) indicated that factor analysis can be made on the data. As a result, a four-factor scale was developed: a) student involvement in assessment (I 19, I 20, I 21 and I 22), b) knowledge on assessment of language skills (I 15, I 16, I 17, and I 18), c) use of assessment results (I 8, I 9, I 10, I 11, I 12, I 13 and I 14), and d) assessment method (I 3, I 4, and I 5). The instrument consisted of five parts, including the first part that collected demographic information of the participants. The instrument was found reliable with .86 Cronbach alpha value. All the sub-factorial groups were also found reliable with .81, .91, .81, and .76 alpha values in consecutively.

3.2. Qualitative data collection through an in-service training session

Subsequently, a full-day INSET session was planned with 20 EFL team leaders. The theme of the session was announced as 'the assessment factor in language teaching' to the invited participants. The training was designed by the researchers to include four parts: a) the significance of assessment in language instruction, b) formative vs summative assessment in language teaching and their uses; c) common assessment terminology and their definitions; d) practical recommendations on assessment and reflection. The researchers of this study delivered the training alternatively. While one trained the participants, the other took field notes and observed.

The first part of the training served as a warm-up and a general introduction. Here, trainers highlighted the importance of assessment in instructional planning in general and in language instruction in particular. In the second part of the training, where the first portion of the qualitative data was collected, participants were asked to indicate whether they knew the definitions of 10 assessment related terminology: assessment, evaluation, testing, measurement, examination, benchmark, outcomes, grading, scoring and washback. There were two reasons for the selection of these terms: First, they were selected to be used during the training by the researchers and the activity served as identification of background information. Second, they were the most frequently referred ones within their context regarding the assessments (international and teacher made) they were using. Each of these terms was then explained and exemplified by the researchers upon collection of teachers' definitions.

For data collection, initially, they were presented a handout consisting of a list of the mentioned terms. They were instructed to write a definition for each term using one of the three different methods: a) write the definition using your existing knowledge (your definition), b) write the definition by using the internet, c) ask for the definition from another participant in the training. Each instruction was given step by step. The participants were given time to write a definition to the terms while both researchers observed the process. They were given a different color pen at each step. Color-coding enabled the researchers to control the process and facilitated the analysis. They were also asked to specify whether they had their own definition (draw 1 star), whether they googled it or asked someone else in the group (draw 2 stars), or they couldn't come up with a definition (draw three stars). The activity sheets were then collected for later analyses. Next, trainers provided correct definitions for each term, participants took notes, and the group compared their definitions to the given ones. These were also checked by the researchers during data analysis to evaluate the correctness of the participants' self-evaluation by using the same definitions that were shared and discussed during the training. At this time, as scientific definitions were not expected, the control of the definitions were made to the extent they reflect the meaning. Lastly, trainers conducted a group discussion for in-depth understanding of the mentioned assessment terminology with real-life examples in order to ensure that all trainers gained a deeper insight about assessment.

In the third part of the training, building on the definitions provided, the researchers started a further discussion on formative and summative assessment. The participants were given another activity handout, in which they were asked to compare formative to summative assessment according to the following domains: relation to instruction, frequency, relation to grading, students' role, requirements for use, purpose, and risk factor. Following a lengthy

discussion on each domain, the researchers provided answers for each domain and gave detailed examples. Lastly, as the second portion of the qualitative data collection process, participant teachers were asked to explain formative and summative assessment (and how they relate to each other) by using a metaphor. The participants were strictly instructed to come up with metaphors (not definitions) to explain the two concepts. The papers were then collected for data analysis. This part of the training ended with researchers sharing participants' metaphors one by one and reflecting on them as a whole group. These two terms (formative and summative assessments) were not involved in the terminology list because they were dealt with throughout the training in different ways by using various techniques. Besides, the researchers aimed at identifying participants' understanding at a higher cognitive level. Both the first activity and the one that involved writing metaphors served to demonstrate knowledge and understanding at different levels.

Finally, the fourth and the last part of the training focused on practical suggestions, reflection on the training as a whole, summary and wrap-up. All participants agreed that the training increased their awareness, knowledge and understanding on assessment.

3.3. Data Analyses

Quantitative data that came from the questionnaire were analyzed via SPSS 23 software program. The data were found to be normally distributed as the skewness and curtosis values were between -1,5 and +1,5 (Tabachnick & Fidell, 2013). Accordingly, dual and multiple comparisons were made by using parametric tests: independent samples t-test and One-Way ANOVA. With regards to the qualitative data, the participants' definitions of assessment terminology were first analyzed by identifying the number of instances that they think they knew the definition of a term. The purpose was to find out the ELT teachers' self-perceptions and knowledge on the issue. Next, their definitions were marked as correct or incorrect by two researchers according to the given definitions during the training. The researcher agreement was 100% in all responses. Finally, participants' metaphors on formative and summative assessment were analyzed. The metaphors were initially categorized according to the domain to which they referred, such as sports, cooking etc. The researchers then created a list of characteristics that were representative of that domain, e.g. listing out the features/steps/abilities required for cooking. Finally, the participant metaphors categorized under a certain domain were compared with the characteristics of formative and/or summative assessment. Accordingly, it was aimed to pinpoint the participants' understanding of the aforementioned assessment types.

4. RESULTS

In order to identify the ELT teachers' perceptions, use and understanding of assessment and assessment terminology, quantitative data were collected via a perception questionnaire. According to the results, the participants have moderately high positive perceptions regarding the four sub-groups of the questionnaire (student involvement in the assessment process M= 3,60; self-perceptions and knowledge on assessing language skills M= 4,14; use of assessment results M= 3,81; and use of appropriate assessment method M= 3.48). When

the factor means are examined, it can be seen that the highest positive self-perception is on their knowledge on the assessment of four language skills. The lowest perceptions, on the other hand, are related to the use of appropriate assessment method. For a detailed analysis, means for participants' responses to each item are given in Table 1.

Table 1. Items' descriptive statistics (N=249).

ITEMS	MEAN	SD
I conduct both formative and summative assessment in my classes.	3,79	1,10
I prepare rubrics to assess language skills.	3,32	1,14
I prepare rubrics to assess task performance.	3,32	1,14
I use assessment results to decide students' success and failure.	3,13	1,20
I use assessment results to give feedback to my students.	3,85	1,11
I use assessment results as feedback about my teaching.	3,86	1,07
I use assessment results to give feedback to parents about their children's development.	3,79	,98
I use assessment results to give feedback to school administration.	3,76	1,02
My colleagues and I (as a team) use assessment results when preparing our syllabus.	3,41	1,08
I discuss assessment results with other colleagues to plan our lessons.	3,78	1,01
I know how to assess the listening skill.	4,09	,74
I know how to assess the reading skill.	4,18	,69
I know how to assess the speaking skill.	4,20	,72
I know how to assess the writing skill.	4,08	,78
I involve my students in the assessment process by asking their assessment preferences.	3,26	1,16
I create opportunities for my students to do self-assessment.	3,85	,93
I implement peer-assessment.	3,64	1,00
I take my students' opinions into consideration in terms of their assessment preferences.	3,63	1,16

Second research question (RQ) aimed to inquire about any possible differences between participants' perceptions in terms of their gender and teaching experiences. According to the t-test and One-way ANOVA results, there were no statistically significant differences in all factorial groups in terms of gender and working experience.

Qualitative data were collected to answer the third, the fourth and the fifth RQs. In order to find out how participants self-evaluate their knowledge and understanding of assessment

terminology, their definitions were examined by both researchers of the present study. The teachers were given a list of 10 assessment terms. Of the 20 teachers, 17 returned the hand-out sheets, 3 teachers did not want to hand in their papers. Totally, there were 170 given definitions. Out of these 170 definitions, 70 were indicated as participants' own definitions, 98 were given either by searching on the internet or asking a colleague. Finally, only two had no response. Later, the terms for which the participants gave their own definitions were tallied and checked for correctness (Table 2). Out of the 70 definitions that participants indicated as their own, only 18 (%25,7) were actually correct.

When the mostly searched or asked terminology was examined, it was found out that washback (n=14), benchmark (n=12), and measurement (n=12) were inquired more than the others. Interestingly examination (n=10) and evaluation (n=10) were also among the ones that were searched by the participants. Overall, the results point out that the participants are in fact not knowledgeable about assessment terminology, even about the ones they think they know.

Table 2. The terms for which participants gave their own definitions.

Term	N	Correct Definitions (n)
Assessment	8	3
Evaluation	7	1
Testing	8	-
Measurement	5	-
Examination	7	2
Benchmark	5	2
Outcome	11	3
Grading	8	5
Scoring	9	2
Benchmark	2	-
Total	70	18

Finally, for the last RQ, the ELT teachers were asked to use a metaphor to define formative and summative assessment and indicate how they relate to each other. First, metaphors given by the participants were categorized (Table 3), then the characteristics of each metaphor group were listed by the researchers to associate these with the features of formative and summative assessment. By this, the researchers aimed to identify the appropriacy of the metaphors used as well as to obtain information regarding the teachers' understanding of the concepts.

Table 3. Metaphors used by the participants.

CATEGORY	METAPHORS	EXAMPLE MEANING UNIT
Sports	Car race Football game Marathon Sailing	Sail on a formative wind and look back from the summative part to the distance of how far you sailed.
Course of life	Youth/old age Birth/death Seed/plant Bird learning to fly Working throughout a year/experience	Formative assessment is a seed that you bury. You water it, it grows and then blossoms. However, summative assessment is the tree (plant) itself that you bought or planted.
Food preparation	Cooking Making a cake	Formative assessment is like cooking. You get better at making dishes and get on going feedback from your loved ones. A summative assessment is more like a dinner party. Your dish must be delicious and you only have one chance to impress your guests.

With a closer look at the metaphors and the table, it can be noticed that the teachers were able to identify the features of formative and summative assessment using appropriate metaphors. All the metaphors provided by the participants focused on the process and the product aspects of the assessment process. The characteristics of these categories are listed in Table 4. Analysis of the metaphors show that the participants have an accurate understanding of formative and summative assessments and how they relate to each other.

Table 4. Features of metaphor categories and assessment types.

CATEGORY	FEATURES	ASSESSMENT TYPE	FEATURES
Sports + Food Preparation	Structured Systematic Linear Step-by-step Skills-based Performance-based	Formative	Process-oriented Continuous Structured Systematic Skills-based Performance-based
Course of life	Effort Process Beginning-end	Summative	Product-oriented Focus on an end-goal of a process

5. Discussion

Results showed moderately high perceptions of the group under investigation. Participants feel comfortable about assessing language skills. However, they seem dubious about how to involve students into the assessment process. They also seem to have a degree of uncertainty in selecting and using the appropriate assessment methods (summative and/or formative). This result indicates participants' limited knowledge in effective use of assessment. In other words, the participant teachers think they can do assessment OF learning, whereas they are not sure whether they can use assessment FOR learning purposes. From this perspective, it can be inferred that assessment in the participants' classrooms has an evaluative purpose and its use to provide feedback to support learning is limited. Teachers' self-perceptions on involving students into the process of assessment is only moderately high, which may indicate a need for support in planning and using assessment in a learning-centered manner. Moreover, the findings indirectly reveal that knowledge about a phenomenon can never ensure its appropriate application or implementation. Hence, supporting pre- and in-service training of teachers via practical elements of assessment seems more than necessary to facilitate the transition of theory into practice.

No differences could be found between genders and groups with different teaching experiences. The possible reason for this might be due to the fact that the number of participants in the groups were not equal. Actually, the female participants strongly outnumbered the male ones. The study focused on teachers at a single institution; hence, it was not possible to change or equalize the number of groups. Although this may also indicate that English teaching profession in Turkish private schools is quite a female-dominant field, obtaining conclusive results in this case might have been hampered by such an imbalance in the participant group. Similarly, sub-groups of participants according to teaching experience were not evenly distributed. Still, no significant difference among groups in this respect shows that participant perception on assessment is not dependent on experience in the field. One might expect that as the experience in teaching increases, so should the perceptions of and awareness about the sub-domains in the profession, such as assessment. However, the results did not reveal such a case. This may also be interpreted as the indication that teachers fail to broaden their expertise and repertoire of skills on assessment along with their years of experience, adhering mostly to traditional summative purposes. Whether a recent graduate from the department or a teacher who graduated more than 12 years ago, most teachers seem to stick to conventional ways when it comes to using assessment as a part of their instruction. This might have several reasons. They either may not be so open to change or they may have not received sufficient training on using different assessment methods. Also, exam-oriented curriculum that has long been implemented in Turkey as well as parental/administrational demands might be other reasons behind these preferences. Considering the positive feedback obtained at the end of the training session from 20 participants, the latter two seem to be within the bounds of a stronger possibility.

Self-reports on assessment terminology show that the teachers mostly gave incorrect definitions to the terms asked in the training. More dramatically, the participants indicated that they considered these definitions true since they did not google them or ask them to another colleague. In other words, the definitions that participants assumed true were considerably incorrect. This reveals that the teachers either have limited understanding of the

assessment terminology or they lack that knowledge. This is contradictory to the quantitative data results as they indicated teachers' moderately high self-perceptions. Both groups of data rely on participants' self-understanding of the phenomena investigated. Hence, cross-check of their understanding proves that they have higher self-perceptions regarding their knowledge on assessment, but they displayed limited knowledge of the basic terminology. The inconsistency of the quantitative and qualitative findings creates a question so as to the extent to which the participants' self-perceptions reflect their actual performance. In other words, perceptions of knowledge cannot be perceived as an actual indicator of understanding unless one shows the ability to display it in novel situations.

The second group of qualitative data consisted of the metaphors given by the participants to explain summative and formative assessment. Data analysis shows that the participants have an accurate understanding of different assessment types. As this group of data was collected subsequent to the part of the training that focused on comparing summative and formative assessment, it can be inferred that the support they received via training had a positive impact on their understanding of assessment methods. It can also be argued that the correct use of metaphors is a better indicator of understanding of concepts when compared to mere definitions. As an indication of higher order cognitive skills, using metaphors demonstrates self-expression and synthesis of the phenomena. This result suggests that engaging teachers with tasks that require higher order thinking on an issue actually has wider and in-depth benefits that contributes to their comprehension. Therefore, the findings emphasize the importance of an effective, skills-oriented training during which the teachers are challenged by activities that require critical thinking.

The research results have confirmed earlier literature as the participants of the present study also displayed inconsistencies of self-perceptions and their actual knowledge and practices (Hatipoğlu, 2015; Mede & Atay, 2017; Ölmezler-Öztürk & Aydın, 2019; Öz & Atay, 2017). The results have also revealed similar findings to previous research that underlines teachers' need to improve their knowledge, skills, expertise and self-confidence at integrating assessment practices into their teaching (DeLuca, Valiquette, Coombs, LaPointe-McEwan, & Luhanga, 2018; DeLuca, LaPointe-McEwan & Luhanga, 2016; Gotch & French, 2014). However, since no significant difference has been found between participant groups in terms of expertise, the results of the current study have not specified this need particularly to novice teachers as some previous studies concluded earlier (Owen, 2016; Popham, 2009; Yamtim & Wongwanich, 2014). Finally, the findings also point out a need for a closer look at pre- and in-service training of teachers on assessment procedures. The results of the current study are in line with the previous research that claim an insufficiency regarding the practical aspect of teacher training in terms of assessment procedures and their uses (Balbay, et.al., 2018; Demirtaş, 2008).

6. Conclusion and Implications

Assessment, being one of the most important components of teaching-learning process, is a rarely investigated issue in terms of its use. Thus, the present study aimed at shedding some light on the issue by investigating ELT teachers' perceptions via an MMR design. The quantitative findings provided some descriptive information regarding teachers'

self-perceptions and provided an insight about their understanding of assessment and use of assessment results. Qualitative data, on the other hand, served two different purposes. Firstly, teachers' self-reports on their knowledge about the assessment terminology together with the evaluation of their definitions revealed valuable information regarding the teachers' actual knowledge of assessment terminology. Secondly, the metaphors they came up with to explain formative and summative assessments following the training provided information regarding the importance of an effective, skills-based teacher training on teachers' critical understanding. Although the quantitative findings and the first group of qualitative data pointed out the inconsistencies of teachers' perceptions and actual knowledge, their use of metaphors indicated the development of understanding at a higher cognitive level.

The current study is significant as it has dealt with a rarely investigated issue by having multiple perspectives on the issue. The results point out the discrepancy between theory and actual practice as well as the inconsistency between language teachers' self-perceptions and actual knowledge about assessment. Hence, it can be concluded that ELT teachers need support and further training on assessment throughout their university education, during practicum and after their graduation. At the outset, prospective teachers at ELT departments commonly gain theory-based knowledge, but the practical aspect of such theory is occasionally neglected (Hudson, Nguyen, & Hudson, 2008; Ölmezler-Öztürk & Aydın, 2019). Knowledge of a concept and its implication require different types of cognitive and practical skills. Thus, further support that aims for the implementation of the theoretical information is more than necessary during pre-service teacher education.

Once teachers start teaching, INSET becomes highly valuable since there is considerable opportunity to combine theory into practice as teachers are actually teaching and forming their instructional understanding. The findings of the current study claim positive effects of INSET that focuses on higher cognitive involvement. The training provided after the quantitative data collection underlines the importance of the content as well as the training techniques and methodology when introducing concepts. When working with practitioners, trainers need to involve some form of higher order cognitive tasks to promote thinking and integrating theory into practice. Using metaphors in teacher trainings as a teaching tool as well as a proof understanding is one of the effective ways in achieving this goal (Erickson, & Pinnegar, 2017; Boyd & Bloxham, 2014).

The research is not without its limitations. The findings are limited to the participant group involved in the study. For the generalizability of the results, further studies are necessary with state school ELT teachers. Moreover, the study is only concerned with self-perceptions. The results showed that perceptions are not always reflected truly by the participants. Hence further studies need to use a different methodology in order to observe teachers performing assessment as a part of their instruction.

7. REFERENCES

Anderson, L.W., & Krathwohl, D.R. (2001). *A taxonomy for teaching, learning, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.

Arslan, D., & Karatas, Z. (2015). Mirror of prospective teachers' mind. Metaphors. *Procedia - Social and Behavioral Sciences, 197*, 1464-1471,

Balagtas, M.U., Dacanay, A. G., Dizon, M.A., & Duque, R.E. (2010). Literacy level on educational assessment of students in a premiere teacher education institution: Basis for a capability building program. *The Assessment Handbook, 4*(1), 1-19.

Balbay, S., Pamuk, G., Temir, T., & Doğan, C. (2018). Issues in pre-service and in-service teacher training programs for university English instructors in Turkey. *Journal of Language and Linguistic Studies, 14*(2), 48-60.

Bloom, B.S. (1956). *Taxonomy of educational objectives, handbook I: The cognitive domain.* David McKay Co Inc.

Boyd, P., & Bloxham, S. (2014). A situative metaphor for teacher learning: The case of university tutors learning to grade student coursework. *British Educational Research Journal, 40*(2), 337-352.

Craig, C. J. (2018). Metaphors of knowing, doing and being: Capturing experience in teaching and teacher education. *Teaching and Teacher Education, 69*, 300-311.

Crusan, D., Plakans, L., & Gebril, A. (2016). Writing assessment literacy: Surveying second language teachers' knowledge, beliefs, and practices. *Assessing Writing, 28,* 43-56.

Demirtaş, Z.T. (2008). *An evaluation of the relationship between teachers' perception of institutional communication and in-service training in primary schools* (Unpublished master's thesis). Yeditepe University, İstanbul, Turkey.

DeLuca, C., LaPointe-McEwan, D., & Luhanga, U. (2016). Teacher assessment literacy: A review of international standards and measures. *Educational Assessment, Evaluation and Accountability, 28*, 251–272.

DeLuca, C., Valiquette, A., Coombs, A., LaPointe-McEwan, D., & Luhanga, U. (2018). Teachers' approaches to classroom assessment: a large-scale survey. *Assessment in Education: Principles, Policy & Practice, 25*(4), 355-375.

Dornyei, Z. (2007). *Research methods in applied linguistics.* Oxford University Press.

Earl, L. (2003). *Assessment as learning.* Thousand Oaks.

Erarslan, A. (2019). Factors affecting the implementation of primary school English language teaching programs in Turkey. *The Journal of Language Teaching and Learning*, 9(2), 7-22.

Erden, H. (2016). Pre-service teachers' perceptions on being a teacher through a metaphoric perspective. *The Anthropologist, 24*(1), 134-147.

Erickson, L. B., & Pinnegar, S. (2017). Consequences of personal teaching metaphors for teacher identity and practice. *Teachers and Teaching, 23*(1), 106-122,

Farley-Ripple, E. N., Jennings, A. S., & Buttram, J. (2019). Toward a framework for classifying teachers' use of assessment data. *AERA Open, 5*(4), 1-18.

Garton, S. (2014). Unresolved issues and new challenges in teaching English to young learners: The case of South Korea. *Current Issues in Language Planning, 15*(2), 201-219. https://doi.org/10.1080/14664208.2014.858657

Giraldo, F. (2018). Language assessment literacy: Implications for language teachers. *Profile: Issues in Teachers' Professional Development, 20*(1), 179-195.

Gürsoy, E., Çelik Korkmaz, Ş. & Damar, E.A. (2017). English Language Teaching within the New Educational Policy of Turkey: Views of Stakeholders. *International Education Studies,(10)*4, 18-30.

Gotch, C. M., & French, B. F. (2014). A systematic review of assessment literacy measures. *Educational Measurement: Issues and Practice, 33*, 14–18.

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.

Herrera, L., & Macías, D. (2015). A call for language assessment literacy in the education and development of teachers of English as a foreign language. *Colombian Applied Linguistics Journal, 17*(2), 302-312.

Hopfenbeck, T.N. (2019). Assessment literacy across contexts and competencies. *Assessment in Education: Principles, Policy & Practice, 26*(2), 119-122.

Izadinia, M. (2016). Student teachers' and mentor teachers' perceptions and expectations of a mentoring relationship: do they match or clash? *Professional Development in Education, 42*(3), 387-402.

Kvasova, O. & Kavytska, T. (2014). The Assessment Competence of University Foreign Language Teachers: A Ukranian Perspective. *CercleS, 4*(1), 159-177.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago University Press.

Leavy, A.M., McSorley, F.A., Boté, L.A. (2007). An examination of what metaphor construction reveals about the evolution of preservice teachers' beliefs about teaching and learning. *Teaching and Teacher Education, 23*(7), 1217-1233.

Low, G. (2008). Metaphor and education. In R. W. Gibbs (Ed.), *The Cambridge handbook of metaphor and thought* (pp. 212-231). Cambridge University Press.

Malone, M. (2013). The essentials of assessment literacy: Contrasts between testers and users. *Language Testing, 30*(3), 329-344.

Mede, E., & Atay, D. (2017). English language teacher's assessment literacy: The Turkish context. *Ankara* Üniversitesi *Tömer Dil Dergisi, 168*(1), 43-60.

Mellado, L., Bermejo, M.L., & Mellado, V. (2012). Personal metaphors of prospective secondary economics and science teachers. *Asia-Pacific Journal of Teacher Education, 40*(4), 395-408.

Miles, M. B., Huberman, A. M., & Saldana, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). Sage.

Nguyen, H. T. M. (2011). Primary English language education policy in Vietnam: Insights from implementation. *Current Issues in Language Planning, 12*(2), 225-249. https://doi.org/10.1080/14664208.2011.597048

Owen, L. (2016). The impact of feedback as formative assessment on student performance. *International Journal of Teaching and Learning in Higher Education, 28*(2), 168-175.

Ölmezler-Öztürk E., & Aydın B. (2019). Investigating language assessment knowledge of EFL teachers. *Hacettepe University Journal of Education, 34*(3), 602-620.

Önalan, O, & Esim, G. (in press). Primary and secondary level EFL teachers' use of assessment and assessment results in Turkey. *Journal on English Language Teaching (JELT)*.

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

Paranosic, N., & Riveros, A. (2017). The metaphorical department head: using metaphors as analytic tools to investigate the role of department head. *International Journal of Leadership in Education, 20*(4), 432-450.

Popham, J.W. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory into Practice, 48*, 4-11.

Popham, W. J. (2013). *Classroom assessment: What teachers need to know* (7th ed.). Pearson.

Saban, A. (2010). Prospective teachers' metaphorical conceptualizations of learner. *Teaching and Teacher Education, 26*(2), 290-305.

Saban, A., Kocbeker, B.N., & Saban, A. (2007). Prospective teachers' conceptions of teaching and learning revealed through metaphor analysis. *Learning and Instruction,17*(2), 123-139.

Sarıçoban, G. (2012). Foreign language education policies in Turkey. *Procedia – Social and Behavioral Sciences, 46*, 2643–2648.

Schmitt, R. (2005). Systematic metaphor analysis as a method of qualitative research. *The Qualitative Report, 10*(2), 358-394.

Selvi, A. F. (2014). The medium-of-instruction debate in Turkey: Oscillating between national ideas and bilingual ideals. *Current Issues in Language Planning, 15*(2), 133-152. https://doi.org/10.1080/14664208.2014.898357

Stoynoff, S., & Chapelle, C.A. (Eds.). (2005). *ESOL tests and testing: A resource for teachers and program administrators.* Alexandria, VA: TESOL.

Tabachnick B.G., & Fidell, L.S. (2013). *Using multivariate statistics* (6th ed.). Pearson.

Treagust, D. F., & Duit, R. (2015). On the Significance of Conceptual Metaphors in Teaching and Learning Science: Commentary on Lancor; Niebert and Gropengiesser; and Fuchs. *International Journal of Science Education, 37*, 958-965.

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.

Veneziano L. & Hooper, J. (1997). A method for quantifying content validity of health-related questionnaires. *American Journal of Health Behavior, 21*(1), 67-70.

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

Wormeli, R. (2009). *Metaphors and analogies: Power tools for teaching any subject.* Stenhouse Publishers.

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

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

Xu, Y., & Brown, G. T. L. (2017). University English teacher assessment literacy: A survey-test report from China. *Papers in Language Testing and Assessment, 6*(1), 149-162.

Yob, I. M. (2003). Thinking constructively with metaphors. *Studies in Philosophy and Education, 22*, 127-138.

Zhao, H., Coombs, S., & Zhou, X. (2010). Developing professional knowledge about teachers through metaphor research: Facilitating a process of change. *Teacher Development, 14*, 381-395.