



HACETTEPE ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Department of Foreign Language Education

English Language Teaching Program

LANGUAGE ASSESSMENT LITERACY LEVEL AND NEEDS OF IN-SERVICE EFL
TEACHERS IN TURKEY

Gamze SARIYILDIZ CANLI

Ph.D. Dissertation

Ankara, 2023

With leadership, research, innovation, high quality education and change,

To the leading edge... Toward being the best...



HACETTEPE ÜNİVERSİTESİ EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Department of Foreign Language Education

English Language Teaching Program

LANGUAGE ASSESSMENT LITERACY LEVEL AND NEEDS OF IN-SERVICE EFL
TEACHERS IN TURKEY

TÜRKİYE'DEKİ HİZMETİÇİ İNGİLİZCE ÖĞRETMENLERİNİN DİL DEĞERLENDİRME
OKURYAZARLIK DÜZEYİ VE İHTİYAÇLARI

Gamze SARIYILDIZ CANLI

Ph.D. Dissertation

Ankara, 2023

Acceptance and Approval

To the Graduate School of Educational Sciences,

This dissertation, prepared by **GAMZE SARIYILDIZ CANLI** and entitled “Language Assessment Literacy Level and Needs of In-Service EFL Teachers in Turkey” has been approved as a thesis for the Degree of **Ph.D.** in the **Program of English Language Teaching** in the **Department of Foreign Language Education** by the members of the Examining Committee.

Chair	Prof. Dr. İsmail Hakkı MİRİCİ
Member (Supervisor)	Asst. Prof. Dr. İsmail Fırat ALTAY
Member	Prof. Dr. Ayhan Kahraman
Member	Asst. Prof. Dr. Özlem Canaran
Member	Asst. Prof. Dr. Hatice Ergül

This is to certify that this thesis/dissertation has been approved by the aforementioned examining committee members on 06/07/2023 in accordance with the relevant articles of the Rules and Regulations of Hacettepe University Graduate School of Educational Sciences, and was accepted as a **Ph.D. Dissertation** in the **Program of English Language Teaching** by the Board of Directors of the Graduate School of Educational Sciences from/...../.....

Prof. Dr. İsmail Hakkı MİRİCİ

Director of Graduate School of Educational Sciences

Abstract

The current research study aimed to examine the LAL levels and training needs of in-service EFL teachers who work in different formal and non-formal education institutions. To that end, training in LTA in undergraduate ELT programs and certain areas they perceive a need for training were investigated. Besides, it aimed to investigate whether training needs in LTA of them differ based on the type of institution and grade level they teach at, their level of experience and education. Their experience concerning LTA practices in classroom was also explored. The study employed a sequential explanatory mixed method design. Both quantitative (questionnaire) and qualitative (interview) data were collected from EFL teachers from different education settings in Turkey. Quantitative data collected from 300 in-service EFL teachers was analyzed using descriptive and inferential statistics. Subsequently, qualitative data collected from 20 in-service EFL teachers to triangulate and clarify questionnaire findings was analyzed utilizing coding and categorizing procedures. The results indicated that the training that they received in four domains of LTA was not adequate and the need for intermediate level of further training was expressed. The findings of the qualitative data also demonstrated that LTA training is not evaluated as sufficient by most of them and the course in LTA did not address the fundamental concepts of LTA adequately. Furthermore, they view the acquisition of expertise in LTA as essential for improving the overall quality of education. Considering these findings and the importance of LAL for EFL teachers, implications and suggestions were provided.

Keywords: language assessment literacy, language testing and assessment, in-service EFL teachers, English language teacher education programs, in-service training needs

Öz

Bu çalışmanın amacı, farklı örgün ve yaygın eğitim kurumlarında çalışan İngilizce öğretmenlerinin Dil Değerlendirme Okuryazarlığı (DDO) düzeylerini ve onların eğitim ihtiyaçlarını incelemektir. Bu amaçla, İngilizce Öğretmenlerinin lisans eğitim programlarında aldıkları dil ölçme ve değerlendirme eğitimleri ve bu alanda eğitim ihtiyacı gördükleri konular araştırılacaktır. Ayrıca, bu çalışma İngilizce öğretmenlerinin dil ölçme ve değerlendirme alanındaki eğitim ihtiyaçlarının öğretim yaptıkları kurum türü ve sınıf düzeyine, deneyim ve eğitim düzeyine göre farklılık gösterip göstermediğini araştırmayı amaçlamaktadır. İngilizce öğretmenlerinin sınıftaki dil ölçme ve değerlendirme uygulamaları ile ilgili deneyimleri de araştırılacaktır. Bu çalışmada sıralı açıklayıcı karma yöntem tasarımı kullanılmıştır. Hem nicel (anket) hem de nitel (görüşme) verileri, Türkiye'deki farklı eğitim kurumlarında çalışan İngilizce öğretmenlerinden toplanmıştır. 300 hizmet içi İngilizce öğretmeninden toplanan nicel veriler, tanımlayıcı ve çıkarımsal istatistikler kullanılarak analiz edilmiştir. Ardından, anket bulgularını güçlendirmek ve açıklığa kavuşturmak için 20 hizmet içi İngilizce öğretmeninden toplanan nitel veriler, kodlama ve kategorizasyon yöntemleriyle analiz edilmiştir. Sonuçlar, İngilizce öğretmenlerinin dört DDO alanında aldıkları eğitimin yetersiz olduğunu ve orta düzeyde ilave eğitim ihtiyacı olduğunu göstermiştir. Nitel verilerin bulguları da, DDO eğitiminin çoğu tarafından yeterli olarak değerlendirilmediğini ve ölçme ve değerlendirme dersinin temel kavramları yeterince ele almadığını göstermiştir. Ayrıca, onlar, DDO alanında uzmanlığın genel eğitim kalitesini iyileştirmek için önemli olduğunu düşünmektedir. Bu bulgular ve İngilizce öğretmenleri için DDO'nun önemi göz önüne alındığında, uygulamalar ve öneriler sunulmuştur.

Anahtar sözcükler: dil değerlendirme okuryazarlığı, dilde ölçme ve değerlendirme, hizmetiçi İngilizce öğretmenleri, İngilizce öğretmeni eğitim programları, hizmetiçi eğitim ihtiyaçları

Acknowledgements

Throughout the extensive journey of completing this dissertation, I have been fortunate to receive support and guidance from various individuals who have made this process smoother and more fulfilling. Foremost, I would like to express my sincere gratitude to my esteemed supervisor, Asst. Prof. Dr. İsmail Fırat ALTAY, whose guidance and encouragement have been invaluable. He has been by my side at every stage, aiding me in my professional growth with his kind feedback, inspiration and endless help. I am grateful for his contributions.

Furthermore, I would like to extend my utmost appreciation to the members of the jury: Prof. Dr. İsmail Hakkı MİRİCİ, Prof. Dr. Ayhan KAHRAMAN, Asst. Prof. Dr. Özlem CANARAN, and Asst. Prof. Dr. Hatice ERGÜL. It has been a truly enriching and productive experience collaborating with them. Thanks to their valuable insights and directions, the goals of the study have been fulfilled. I am grateful for their contribution and the fruitful interactions we had during this process.

I would like to thank my friends who have been an incredible source of support and inspiration throughout the process of completing this dissertation. Their encouragement, understanding, and camaraderie have made this journey more fulfilling and enjoyable. I am also grateful to my dedicated colleagues who have provided valuable insights and feedback throughout the course of my research. Our discussions, collaborations, and shared knowledge have contributed immensely to the depth and quality of this dissertation. I appreciate the intellectual stimulation and exchange of ideas that we have fostered together.

With profound gratitude, I extend my heartfelt appreciation to my family members who have been unwavering sources of love and support in my life. My mother Şule and my father Fedakar Sarıyıldız have provided me with unconditional care and guidance, and I am forever grateful for their presence. My grandmother, Nevin, has always held the best wishes for my success, and I am thankful for her unwavering support. My sister Seda, has brought

joy and happiness into my world, making every moment more delightful. Thanks for being my confidante and my biggest cheerleader, Seda. Their presence has provided me with a sense of comfort and motivation, and I am fortunate to have them by my side. They have been my pillar of strength, always providing a listening ear, words of wisdom, and a source of inspiration. Their unwavering faith in me has given me the confidence to overcome challenges and pursue my goals. I also want to express my thanks to my husband Alpaslan Canlı. He was very kind and understanding through this process. I want to thank him for believing in me. Words fail to express the depth of my gratitude for their presence and the role they have played in my life. I must emphasize that this achievement would not have been possible without their unwavering support. Their love, encouragement, and belief in my abilities have been the foundation of my journey. Their presence has provided me with strength during challenging times and has given me the motivation to persevere. I am forever indebted to them for their continuous support and the indelible impact they have made on my life.

Last but not least, special thanks to the participants of the study, the in-service EFL teachers working across the country. Without their contribution, insightful remarks and helpful suggestions, the study could not have been completed.

To my adored family and

To my cherished sister

Table of Contents

Acceptance and Approval	ii
Abstract	iii
Öz.....	iv
Acknowledgements	v
List of Tables	ix
List of Figures.....	xi
Symbols and Abbreviations	xii
Chapter 1 Introduction.....	1
Statement of the Problem	3
Aim and Significance of the Study.....	4
Research Questions.....	5
Limitations.....	6
Definitions	7
Chapter 2 Literature Review.....	9
Fundamental Concepts in Assessment.....	9
Assessment Literacy.....	11
Language Assessment Literacy	17
EFL Teacher Education in the World	20
EFL Teacher Education in Turkey.....	20
Studies on LAL of in-service and pre-service EFL Teachers in Turkey.....	33
Studies on LAL of in-service and pre-service EFL Teachers Abroad.....	36
Chapter 3 Methodology	46
Research Design.....	46
Setting and Participants	47
Instruments	53
Data Collection Procedure	56

Data Analysis	62
Chapter 4 Findings and Discussion	66
RQ1: In-Service EFL Teachers' Received Training in Language Testing and Assessment	66
RQ2: EFL Teachers' In-service Training Needs in Language Testing and Assessment	75
RQ3: Differences between EFL Teachers' Training Needs.....	84
Results for RQ 4.....	88
Discussion of the 1st Research Question	114
Discussion of the 2nd Research Question	122
Discussion of the 3rd Research Question	140
Discussion of the 4th Research Question	145
Chapter 5 Conclusion and Suggestions	193
Summary of the Study.....	193
Implications of the Study	200
Limitations of the Study	206
Suggestions for Further Research	207
References	209
APPENDIX-A: LAL Questionnaire	ccxxii
Appendix-B: Semi-Structured Interview Form in Turkish.....	ccxxix
Appendix-C: Semi-Structured Interview Form in English.....	ccxxx
Appendix-D: Interview Schedule	ccxxxix
Appendix-E: List of Codes for the Interview Form	ccxxxix
APPENDIX-F: Ethics Committee Approval.....	ccxxxiv
APPENDIX-G: Declaration of Ethical Conduct	ccxxxv
APPENDIX-H: Dissertation Originality Report	ccxxxvi
APPENDIX-I: Yayımlama ve Fikrî Mülkiyet Hakları Beyanı	ccxxxvii

List of Tables

Table 1	<i>The Structure of ELTE Programs in Turkey</i>	22
Table 2	<i>List of In-service trainings that are available to EFL Teachers in Turkey</i>	26
Table 3	<i>Profile of the Participants</i>	49
Table 4	<i>Summary of the Participants' BA, MA and PhD Degrees</i>	50
Table 5	<i>Summary of the interview participants' background information</i>	52
Table 6	<i>Semi-Structured Interview Questions for the In-service EFL Teachers</i>	55
Table 7	<i>Overview of the data collection tools, the corresponding research questions, data analysis methods, the profile and number of participants</i>	56
Table 8	<i>Questionnaire Parts and Reliability Scores</i>	59
Table 9	<i>Normality Test of the Questionnaire</i>	62
Table 10	<i>Summary of the findings for training received in classroom-focused LTA</i>	67
Table 11	<i>Summary of the findings for training received in Knowledge of Testing and Assessment</i>	69
Table 12	<i>Summary of the findings for training received in Purposes of Testing</i>	70
Table 13	<i>Summary of the findings for training received in Content and Concepts of LTA</i> 71	
Table 14	<i>Summary of the findings for training received in four areas of LTA</i>	74
Table 15	<i>Summary of the findings for in-service training needs in classroom-focused LTA</i>	75
Table 16	<i>Summary of the findings for in-service training needs in Knowledge of Testing and Assessment</i>	77
Table 17	<i>Summary of the findings for in-service training needs in Purposes of Testing</i> ...79	
Table 18	<i>Summary of the findings for in-service training needs in Content and concepts of LTA</i>	80
Table 19	<i>Summary of the findings for in-service training needs in four areas of LTA</i>	83
Table 20	<i>Summary of the ANOVA results for educational settings</i>	85
Table 21	<i>Summary of the ANOVA results for grade level of teaching</i>	85
Table 22	<i>Summary of the ANOVA results for years of experience</i>	86
Table 23	<i>Summary of the ANOVA results for level of education</i>	87

Table 24 *Summary of the Tukey HSD results for level of education*88

List of Figures

Figure 1 <i>A Conceptual Framework of Teacher Assessment Literacy in Practice (Xu, & Brown, 2016)</i>	Error! Bookmark not defined.
Figure 2 <i>Explanatory Sequential Design (Creswell, 2012, p.541)</i> Error! Bookmark not defined.	
Figure 3 <i>Histogram for Normality Test of the Questionnaire</i>	63
Figure 4 <i>Normal Q-Q Plot for Normality Test of the Questionnaire</i>	64
Figure 5 <i>Qualitative Data Analysis Process (Creswell, 2012, p. 237)</i>	65

Symbols and Abbreviations

LTA: Language Testing and Assessment

EFL: English as a Foreign Language

ELT: English Language Teaching

LAL: Language Assessment Literacy

Chapter 1

Introduction

Testing and assessment are of paramount importance to curriculum and education system (Alderson, 2005; Richards, 1984) in that educational practices including teaching and learning processes are closely related to testing and assessment practices (Dochy, 2009). Successful and quality education is possible through qualified testing and assessment procedures and practices; in other words, without proper and well-designed testing and assessment practices, it is impossible for quality education to be delivered. This has been further highlighted with “educational assessment” which refers to the indispensable connection between learning and assessment and using the results of assessment methods and tools for the purpose of enhancing student learning and programs (Stiggins, 2008). Besides the conventional purposes of assessment like demonstrating how many of the learning goals and objectives have been achieved and what learners attained, which is called as Assessment of Learning (AoL), assessment and testing can be conducted to ascertain and meet the needs of students by emphasizing the areas that require further elaboration, which is referred to as Assessment for Learning (AfL). While the former serves the purpose of revealing what is achieved in terms of educational objectives and goals, the latter strives to give constructive feedback to students and scaffold learning aside from refining teaching methods and techniques (Black and William, 1998; Wood, 2007).

Taking the importance of testing and assessment practices for qualified learning and teaching processes into account, it is fundamental for teachers to have adequate knowledge and skills of designing, evaluating, and applying various assessment practices, which is referred to as “assessment literacy”. It can be described as the capability to differentiate between high-quality and poor-quality assessment practices and utilizing this for making inferences related to learners’ success and progress (Stiggins, 1991). Taking this a step further, Popham (2004) states the nonexistence of assessment literacy as “professional suicide” (p.82) to emphasize the crucial role of assessment literacy for teachers as

professionals. In this regard, being assessment literate is highlighted as one of the vital competencies of teachers for achieving quality in teaching practices. In the same vein, Language instructors are required to possess proficiency in language assessment, which is described as “language assessment literacy”.

The need for language teachers to possess sufficient knowledge and education in the domain of language assessment is substantial taking into account the expansion and importance of evaluations and measurements in the area of language. Within this context, language assessment literacy (LAL) is of paramount importance in language education since it allows teachers to grasp and apply information related to students' performance and contributes to the language learning process. In spite of the importance of assessment in language education context, teachers are not given sufficient training in language testing and assessment, hence the majority of them are deficient in terms of knowledge, practice, experience, and confidence when it comes to assessment. (Brookhart, 2001; Taylor, 2009). Therefore, a limited number of teachers are deemed adequate for fulfilling the demands and tackling the difficulties of language assessment. (Stiggins, 2002). Furthermore, language teachers are found to have low assessment self-efficacy owing to the fact that they received either insufficient or no training on testing and assessment (Deluca & Klinger, 2010; Tsagari & Vogt, 2017). Different studies have also revealed that EFL instructors have either low or average level of LAL since their competence in the design of valid, reliable and appropriate assessment tools is insufficient (Haznedar, 2012; Kiomrs, Abdolmehdi, & Naser, 2011; Köksal, 2004; Leaph, Channy, & Chan, 2015; Vogt & Tsagari, 2014).

In the last two decades, researchers' attention has been directed to the issue of language teachers' testing and assessment knowledge and practices with regard to the problems and insufficient LAL level of teachers (Hatipoğlu, 2010; Jeong, 2013; Stiggins, 1999, 2001; Vogt & Tsagari, 2014). Nonetheless, research on LAL and the course-based language testing and assessment training provided to pre-service and in-service ELT teachers are not enough considering its pivotal role in language learning and teaching

processes. Language teachers' preparation regarding language assessment prior to their entry to their professions plays a significant role in enhancing language assessment literacy and improving their professional identity as well as the quality of language assessment and teaching practices (Darling-Hammond, 2000). Notwithstanding the importance of language assessment literacy in foreign language teacher education, there are few studies which have been conducted to investigate the LAL training of pre-service language teachers in Turkey. As for in-service language teachers who are responsible for classroom-based assessment along with standardized language tests in their professions, limited research has been carried out to investigate their level of LAL and perceived LAL training needs in Turkey (Mede & Atay, 2017).

Upon revision of the studies in this field, it is apparent that further investigation is necessary to determine the LAL levels and training needs of English language teachers who are currently employed in diverse educational environments in Turkey. Therefore, this study aims to explore LAL levels and training needs of in-service EFL teachers and to bridge this crucial gap in the field. It also aims to investigate whether in-service English language teachers' training needs differ from one another in terms of their level of experience and education. Moreover, this study seeks to find out whether training needs of in-service EFL teachers differ by the grade level they teach and different educational settings. Furthermore, it strives to find out language teachers' descriptions of the experience in LTA practices as teachers in classroom contexts. All in all, this study aims to examine the development of in-service English language teachers' course-based LAL and their further training needs and experience in LTA field through the analysis of different dimensions (i.e., different educational settings, the grade level they teach, level of education and experience).

Statement of the Problem

The main problem of this study is to investigate the LAL and training needs of in-service EFL teachers regarding language testing and assessment (LTA). The current level

of LAL and training needs in LTA field of in-service EFL teachers are aimed to be examined. More specifically, to what extent in-service EFL teachers are trained in the language assessment field, and they need training in this field will be explored. In addition, whether in-service EFL teachers' training needs differ from one another regarding their education and experience level. Furthermore, it strives to examine whether there is a significant difference in the training needs of in-service ELT teachers by the grade level they are teaching at the present and different education settings. Finally, it aims to explore the experience of language teachers regarding their LTA practices in classrooms.

Aim and Significance of the Study

This current study seeks to examine the status quo of LAL and training needs of in-service EFL teachers who work in formal and non-formal education institutions. To this end, this study seeks to delve into LAL levels and further training needs of in-service EFL teachers. Therefore, to which extent they had training in LTA in their teacher education programs and to what extent they believe they need training in LTA areas will be investigated. Also, it aims to explore whether there is an important difference in the pronounced needs of training in areas of LTA based on their education and experience level, the grade level they are teaching at the present and different education settings. Another aim of this study is to find out the experience of language teachers in terms of LTA practices in their teaching contexts.

The importance of the research can be identified from multiple perspectives. Firstly, given the prominence of assessment in language education, language teachers should possess LAL in their undergraduate education and foster it in their future professional lives. In this regard, examining the beliefs of in-service language teachers with regard to whether the training in language assessment that they receive in teacher education programs is adequate and in which fields they need training is compulsory in that it gives insight into the status quo of their LAL levels and further training needs. Secondly, investigating the level

of LAL and perceived LAL training needs of in-service language teachers who are directly influenced by the results of assessment training as one of the most important key stakeholders is critical since it may inform in-service assessment courses/programs in terms of revealing the further training needs of teachers in language assessment. In other words, the results can serve as a needs analysis for further decision making in in-service programs. Also, despite the fact that the current study has been conducted in Turkey, the results of this study can serve as a contribution to other English language teacher education programs since it presents a deeper understanding of the importance of LAL for English language teachers. Drawing on the results of the study, tailor-made assessment courses catering for the needs of the teachers can be designed so that effective and qualified assessment practices can be conducted. Moreover, language teachers will find the opportunity to reflect on their knowledge, practices and experience in LTA as well as revising LTA areas that are problematic in their current practices. Given that reflection in teacher education is one of the major areas that is underlined and being a reflective practitioner is a way to improve professional competence and knowledge, language teachers' reflections on LTA field will be of paramount importance to their continuing professional development. More specifically, through such thoughtful reflections they can be more motivated to take assessment courses and improve their competence in LTA, which will enhance their autonomy and awareness of LAL construct, too. Lastly, considering the limited number of studies on LAL of in-service language teachers and their training needs both in Turkey and in the world, this study will be important in terms of shedding light on LAL and language assessment fields in which they need training.

Research Questions

The following research questions are aimed to be answered in this study:

1. How do in-service EFL teachers evaluate their undergraduate training in language testing and assessment? Do they perceive this training as adequate?

2. To what extent do in-service EFL teachers perceive a need for in-service training in language testing and assessment?

3. Do in-service EFL teachers' training needs in language testing and assessment vary according to

a. different educational settings?

b. the current level of education that they are teaching?

c. year of experience?

d. their education level?

4. How do in-service EFL teachers describe their experience in language testing and assessment practices in classroom context?

Limitations

The major limitation of the present research study is the sample of the study in that the data was collected from in-service EFL teachers who work in formal and non-formal education institutions. Even though there was a large number of participants, it did not involve all in-service ELT teachers in Turkey. Therefore, the number of the participants can still be increased to have a good grasp of the LAL and language assessment training of EFL teachers and to generalize the results for all in-service EFL teachers in Turkey. Thus, the findings can be generalized only to that sample group and further studies on other groups are suggested.

Another limitation is that this study solely concentrated on the participation of EFL teachers without involving other crucial stakeholders such as policy makers, program designers, school administrators, teacher trainers, Education faculty members, pre-service teachers and language learners. Considering the perspectives of these stakeholders would provide valuable insights for shaping in-service training programs. Nonetheless, the findings

obtained from this study can still serve as a foundation for the development of tailored in-service training programs in LTA specifically designed for EFL teachers in Turkey.

Also, in this study, a mixed-methods approach was employed, combining quantitative and qualitative research designs. To gain a comprehensive understanding of teachers' language assessment literacy levels and identify their training needs in Language Testing and Assessment (LTA), data were collected through questionnaires and interviews. However, it is worth considering the inclusion of additional data sources such as classroom observations or analysis of teachers' assessment samples. The integration of multiple data sources allows for a more comprehensive exploration of teachers' assessment practices, providing a richer understanding of their capabilities and areas that may require further training.

Definitions

Assessment: A generic term that points out a continuous procedure of eliciting, evaluating and making comments on students' performance and informing learning progress besides identifying students' strong and weak points (Boston, 2002; Brown, 2004; Taras, 2005).

Assessment Literacy: Having a grasp of principles, various functions and aims of assessment, and putting them into practice in different contexts as well as interpreting the results of these assessments accurately and using these results to enhance learning (Popham, 2004; Volante & Fazio, 2007).

Language Assessment Literacy: Having knowledge and skills for creating and evaluating language tests and being familiar with theories, principles and procedures of assessment as well as being aware of the importance of assessment for people and foundations (Fulcher, 2012).

Testing: Measuring the knowledge, performance, or ability of learners at a specific point of time when learners are aware of the evaluation and measurement taking place (Brown, 2004).

Inservice Teachers: Teachers who already have certification and currently teach in the classroom.

Chapter 2

Literature Review

Fundamental Concepts in Assessment

Assessment, testing, measurement, and evaluation are significant components in language education. As these concepts are used interchangeably, it is crucial to differentiate among these terms.

To start with the term “assessment”, it is used in many ways in educational sciences and there is not an agreement regarding what it means accurately (Bachman, 2014). Various terms comprising “testing”, “evaluation”, and “measurement” are utilized alternatively to express “assessment”. According to Lambert and Lines (2000), assessment is the process in which information about students’ replies to educational exercises are collected, commented and noted down. As for Purpura (2016), it expresses “a systematic procedure for eliciting test and non-test data for the purpose of making inferences or claims about certain language-related characteristics of an individual” (p.191). Similarly, Brown (2003) defines it as a continuing process of collecting information about students’ performance either consciously or unconsciously. Through assessment, teachers and educators get beneficial and instant feedback on what and how much students have learned on a given subject. Bachman (2014) gives two characteristics of “assessment” which are systematic and substantively grounded, which differentiates assessment from informal observation. The former refers to the explicit design and application of assessment and is related to the reliability principle as well as providing others to regenerate it in case of need. The latter refers to laying the foundation for commentary of quantitative and qualitative findings of assessment in that it should be a broadly accepted theory about the nature of language usage, learning, skill, prior studies and practice, and it is related to validity principle. As Clapham (2000) put forward, assessment can be stated as a generic term

involving many strategies and methods like tests, observation and interviews to inform teaching and learning processes.

As for “testing”, it is defined as “a set of tasks or activities intended to elicit samples of performance which can be marked or evaluated to provide feedback on a test taker’s ability or knowledge” (Coombe, 2018, p. 41). As the definition reveals, certain tasks are utilized to reveal particular instances of performance associated with some features and skills that cannot be observed. In a similar vein, it is referred to as a method of gauging people’s skill or knowledge in a specific area (Brown, 2004). According to Lynch (2003), testing is more related to quantitative measurement of language skills and objectivity as well as focusing on validity. Therefore, testing can be stated to be a way or method of measurement of a person’s knowledge, skill, performance using tests.

“Measurement” is defined as “the process of quantifying the characteristics of an object of interest according to explicit rules and procedures” (Bachman, 2005, p.8). In other words, it is a kind of assessment including quantification. The features and skills such as vocabulary range and language talent that are aimed to be gauged cannot be observed straightforwardly. This is realized by means of reliable and valid tests, well-defined test specifications and criteria as well as administration procedures, which connects unobservable feature being gauged to observable performance being evaluated.

Another crucial concept that is related to assessment is “evaluation” which means interpreting the findings of assessment procedures and making decisions. It is described as an umbrella term and systematized collection of information for deciding and supporting learning. Four levels of evaluation are listed, which are learner feedback, learner learning, learner behavior, and learning results (Coombe, 2018).

All in all, testing, assessment, measurement and evaluation are all important components in education programs. Testing is more specific than assessment and gauges students’ accomplishment through asking learners to carry out a particular task that is later marked to reveal language skill. Therefore, it places more emphasis on measuring language

skill quantitatively. Assessment, on the other hand, comprises language tests as well as alternative assessment like self-assessment, instructor observation and portfolios which do not require the strict restrictions of tests (Nagai, 2020). While evaluation involves appraising not only the learners but also all features of quality of language program through information obtained from various sources, measurement is related to the quantification of observed attributes of learners (Coombe, Folse & Hubley, 2007).

Assessment Literacy

Assessment literacy, which is fundamental for the success and quality of teaching and learning, was first coined by Stiggins (1991) and described as the capability to recognize the difference between low-quality and high-quality assessments and utilize that to make a conclusion regarding learner accomplishment. In other words, assessment literacy is described as having a grasp of how to assess learners' knowledge and progress and evaluate them to improve their own teaching and learners' development (Webb, 2002). Moreover, Inbar-Lourie (2008) stated that assessment literacy entails the capability to examine the objective and the outcomes of assessment.

Furthermore, Popham (2004) asserted that inadequate competence in testing and assessment; in other words, illiteracy in assessment was described as "professional suicide" (p.82). In this regard, teachers should be assessment literate that will enable them to conduct sound assessment practices for the enhancement of learning opportunities and professional developments. According to Newfields (2006), the significance of being assessment literate can be explained from three regards. In the first place, when time and money allocated for assessment is taken into consideration, which is a considerable amount, it is clear that assessment is a key component of education. In the second place, assessment literacy provides the ability to deduce the findings of literature in the education field; therefore, absence of literacy in assessment poses a challenge to draw conclusions from a variety of academic papers and research studies related to their majors. Lastly,

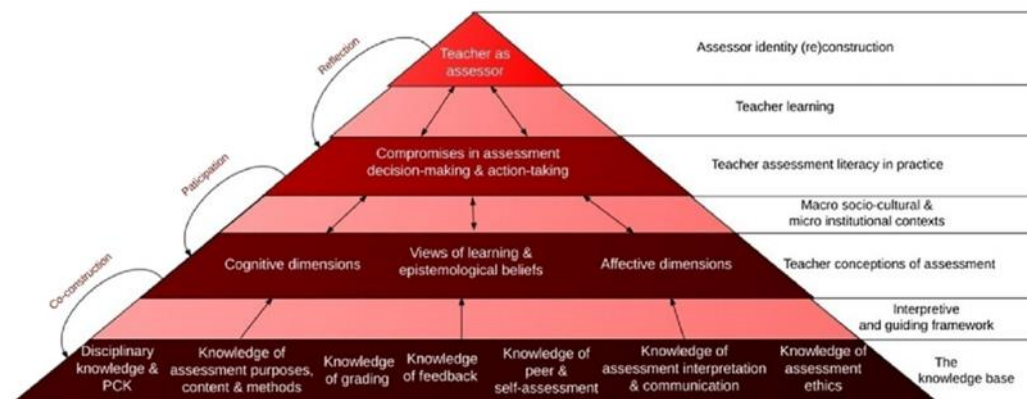
teachers can reveal the outcomes of their own classrooms with their colleagues, which enhances lifelong learning skills.

A growing number of studies have revealed that the level of assessment literacy among pre- and in-service teachers is low and teachers receive limited training in assessment (Alkharusi, 2011; Brookhart, 2001; Malone, 2013; Mertler, 2009). The scrutiny of the relevant literature put forward that both pre- and in-service teacher training programs are proposed as ways to tackle problems in the levels of assessment literacy (Mertler, 2004; Popham, 2009; Taylor, 2009). As a result, a great deal of research on assessment literacy has highlighted the significance of taking the needs of prospective teachers and in-service teachers regarding assessment into consideration and the essential role of teacher education programs with regard to providing them with sound foundations in assessment (Hasselgreen, Carlsen, & Helness, 2004; Herrera & Macías, 2015; Inbar-Lourie, 2008; Mede & Atay, 2017; Popham, 2009).

Xu and Brown (2016) put forward a framework of TALIP after reviewing relevant studies and literature as the figure below shows. It is stated that previous frameworks like that of DeLuca (2012) are helpful for giving insights and a new framework is significant for connecting teacher education and educational assessment involving both pre- and in-service teacher education broadly. This framework does not only include the review of AL studies conducted before, but also that of teacher education. There are six constituents of this framework which are knowledge base, teacher conceptions of assessment, institutional and socio-cultural contexts, teacher assessment literacy in practice (TALIP), teacher learning, and teacher identity (re)construction as assessors.

Figure 1

A Conceptual Framework of Teacher Assessment Literacy in Practice (Xu, & Brown, 2016)



Knowledge base which is the baseline for all other constituents is at the bottom of this framework. As Maclellan (2004) expressed, for sound practices in assessment all teachers need this knowledge. Considering these practices, a good education system requires systematicity and formality regarding assessment. Though it is crucial, it is stated to be insufficient for TALIP for some reasons like not being prepared solutions to problems happening in a complicated classroom setting. Under this component, there are 7 subcomponents that are disciplinary knowledge & PCK, knowledge of assessment purposes, content & methods, knowledge of grading, knowledge of peer & self-assessment, knowledge of assessment interpretation & communication and knowledge of assessment ethics. The first subcomponent is related to the content to be taught and disciplines as measurement of curriculum content is related to educational assessment. The second one comprises different issues like the necessity for teachers to be aware of why they make assessment and the relation of methods of assessment and learning targets and content, diverse strategies of assessment. Knowledge of grading refers to the knowledge that teachers should have and includes logic behind grading, consistency, criteria for rubrics and giving grades. The fourth subcomponent is feedback knowledge which states that

teachers should know various types of feedback in terms of their aims, principles, roles, strength, and shortcomings of them. The fifth one is the awareness of importance of engaging students in assessment such as self and peer-assessment) and ways to utilize these effectively in the class context. The sixth one is related to the necessity for teachers how to interpret the evidence from assessment along with the knowledge of sharing the result to different groups like learners, teachers, and managers. The last one regarding assessment ethics, it is expected that teachers should grasp the liabilities such as the usage, archiving and announcing the results (Tierney, 2013). That component also comprises non-discrimination, and justice. This first level of the framework is dynamic and functions as a base and those passing it will be involved in assessment more profoundly.

The second layer of this framework is teacher conception of assessment as an interpretive and guiding frame. Given that knowledge base is formed through many channels like lectures and seminars, what is obtained is screened out and interpreted by conceptions of teachers. In this regard, their conception functions as a mediator between theories and application of assessment. What teacher conceptions include is not limited to cognitive points, but also affective points. The former points out teachers' opinions regarding true and false about assessment, and it is highlighted that there is a tendency to utilize the knowledge, methods of assessment that are compatible with grasp of teachers while refusing others that are not harmonious. The latter renders emotional features that teachers have already had like their experience of positive and negative feelings towards assessment (Crossman, 2007). These emotional factors are affected by individual and social factors as well as personal and educational practices. Teachers' existing perceptions of assessment could be altered in the long term to boost assessment literacy of teachers.

The third one is micro- and macro- contexts, which impacts teachers' practices of assessment in many ways. It could be both narrow and wide scopes like there can be limits for teachers for what can be done or what should not be done according to policies. Similarly, what teacher candidates learn about assessment is limited by local and national

policies influencing assessment training structures, which can either promote or hinder development of assessment literacy. Also, it is denoted that prospective teachers' learning about assessment during observations in the practicum is affected by the assessment that is conducted by mentor teachers in schools. As to in-service teachers, their practices in assessment are also affected by several stakeholders involving parents, learners, workmates, and administrators (Xu & Liu, 2009).

The next component of this framework is TALIP. It is mentioned that decision making of teachers is the process in which both external dimensions and limits and internal dimensions like beliefs of teachers are balanced. This reveals that for assessment to be good, it should involve reunification of these two parts as Carless (2011) uttered. This assessment literacy in practice is compromised between macro and micro contexts and conceptions of assessment of teachers. Sometimes, external things affect that in terms of the size of the class or curriculum. To give an example, though a teacher finds giving performance assessment logical, because of the population of the class, s/he can have to apply paper and pencil exams. TALIP is highlighted not to be static, but rather dynamic and complicated and represents the truths in the real classroom contexts (Xu & Brown, 2016).

For the fifth component, it is teacher learning and as being assessment literate is a dynamic and ever-growing process, teacher learning should be ongoing, too. If they oftentimes do the same thing for assessment and experience no conflict because both external factors and their own beliefs are consistent, which would result in traditional practices which is not suggested by research regarding the effectiveness of assessment or present-day policy expectations. With respect to efficiency of assessment in classroom, both internal and external factors are influential such as learners' readiness both mentally and emotionally, teachers' acquaintance with learners, subject matter, and the course. According to Koh (2011), teacher learning can come true both inside the class and outside on the condition that they are conscious of the relation among different processes and mechanisms. Two criteria are expressed to obtain teacher learning and these are attending

community activities and doing reflective practice. The first one states that participating in such activities provide them with professional conversations with workmates and gaining insights besides sharing their own perceptions. The second one is highly stressed in that reflection is a must and it stimulates development of them. It is also repetitive and involves analyzing past experience with future actions to realize, which all contribute to teacher development and learning while they are learning and relearning related to assessment theories and practices. All in all, these two factors are emphasized to improve AL and assessment.

At the top of this framework is the teachers' identity as assessors. Conventionally, teachers' role is considered to be instructor and decisions about learners are made externally. Generally, teachers are supposed to make students ready for the exams to be administered, but with formative assessment policies, they are supposed to assess learners to make decisions in pedagogy and this changes the role of teacher into "assessor". Two roles of teachers as both instructors and assessors of learning pose a challenge in educational assessment. This identity (re)construction is stated to be the final aim of TALIP. In this regard, prospective teachers are said to have certain changes from being student to becoming teachers taking charge of assessing and evaluating students. In-service teachers are emphasized to discuss their roles as assessors while interacting with others. Furthermore, this identity construction is stressed to be an unconditioned result of teacher learning as they engage in reflection and participation. This will make them more aware of the worth and the power of assessment for making positive changes in learning of students. Such a consciousness strengthens teachers by providing more autonomy and chances for AL.

More specifically, it is specified that if teachers see themselves only as instructors, teaching could stop after they deliver the subject matter and support students. However, if they are aware of their "assessor" role, they apply assessment practices more consciously to gauge whether they have learnt the content delivered to them and to give effective

feedback promoting progress of learners. In this context, teachers do not only think that this is directly related to student knowledge or skill, but rather they reflect upon their practices and contemplate about alternative ways to enhance the current practices. Thus, starting to build identity as “assessor” is pivotal to enhance beliefs and practices of assessment.

Language Assessment Literacy

Language assessment literacy has derived from the generic term “assessment literacy” and they have common constituents; however, LAL differs from AL in terms of the understanding of language, usage and pedagogy of language (Brookhart, 2001; Giraldo, 2018). LAL is defined as having a sound grasp of assessment including the knowledge, principles and functions of various assessment methods as well as evaluating the results of these assessment procedures accurately and making precise decisions about learners (Lam, 2015). In this regard, a language teacher who has an adequate level of LAL has sound principles and knowledge of language assessment and various assessment methods so as to grasp, design, examine and interpret her/his assessment procedures in the language classroom (Malone, 2013; Scarino, 2013). Another definition of LAL was put forward by Fulcher (2012) following the research on investigation of training needs of teachers regarding language assessment. As for the components of LAL, he asserted that there are three components, which are having the grasp and ability to create, analyze and interpret different kinds of tests and assessments, being aware of different principles and constructs for effective assessment practices, and having the capability to put theoretical knowledge and abilities into broader social, historical and political contexts as well as analyzing the effect of testing on people, community and several foundations.

On the other hand, Davies (2008) asserted that LAL is composed of three factors that are knowledge, skills and principles. While knowledge is related to the language, evaluation and context, skills are explained with regard to creating language tests and drawing conclusions from the results. The principles refer to employing convenient tests,

justice and impacts of assessment on a variety of subjects like people and institutions. As for Scarino (2013), two components were claimed to compose LAL, which are knowledge and process base. While the former refers to the theories of language field that support assessment practices, the latter indicates progressive stage of assessment in language as analyzing and commenting on the notions of assessment.

Considering the relationship between language assessment and teaching practices, it is of pivotal importance for language teachers to improve their assessment knowledge and skills. Especially in the modern era, language teachers have more responsibilities with more focus on formative assessment; in other words, assessment for learning compared to the past where the main focus was on summative assessment, namely assessment of learning (Csépes, 2014). To put it in a different way, assessment does not only serve as a demonstration of what has been learnt and whether goals have been achieved or not but also as a way to obtain and give constructive feedback to learners as well as evaluating and enhancing their approaches and methods. In this regard, language teachers are regarded as both teachers and evaluators of their own teaching environments, which makes being competent at LTA inevitable for them.

Moreover, as Newsfield (2006) asserted being aware of assessment concepts, following the procedures, evaluating the findings will both contribute to students' achievement and motivation towards the foreign language learning besides helping language teachers to cater to the needs of learners during foreign language learning process. In a similar vein, in addition to scaffolding learning, language teachers' competency in language assessment makes a huge contribution to their professional development (Buyukkarci, 2016). Another reason for the requirement of LAL on the part of language teachers is that new advancements in language teaching entail brand new qualifications by language teachers. Language teachers are expected to keep up with these developments and obtain these qualifications for professional development and quality teaching. One of these areas is related to assessment in that the European Language

Portfolio (ELP) highlights peer- and self-assessment as key components of education, and language teachers' awareness and competency in these key concepts is of paramount importance (Morrow, 2004). Therefore, these have to be included in the pedagogy of language teachers and should be possessed by them to have a higher level of LAL considering the limited training and pronounced need for advanced training in these crucial concepts as expressed in a growing body of research (Hatipoğlu, 2015; Lam, 2015; Vogt & Tsagari, 2014; Volante & Fazio, 2007). In short, supporting learning and promoting teacher development as well as keeping pace with new developments in the field and increasing the positive washback on educational practices is enhanced through LAL.

As mounting evidence reveals, what constitutes much of language teachers' assessment knowledge is the courses in teacher education programs and their experience (Popham, 2009; Siegel & Wissehr, 2011; Taylor, 2009). A growing body of literature demonstrated that level of LAL of pre- and in-service language teachers is not high and training in LTA is limited. Furthermore, underdeveloped LAL of pre- and in-service language teachers was attributed to either lack of formal training in LTA (DeLuca & Johnson, 2017; Guerin, 2010; Lam, 2015; Mertler, 2004; Popham, 2006, 2009; Volante & Fazio, 2007) or in-service training in LTA (Büyükkaracı, 2016; DeLuca & Johnson, 2017; Mede & Atay, 2017; Mertler, 2009). In light of these, assessment courses in both pre-service teacher education programs and in-service training courses are recommended to enhance LAL of both pre- and in-service EFL teachers. While through pre-service courses in LTA prospective teachers' awareness and expertise is developed and shaped, in-service trainings in LTA strengthen their knowledge and competencies as they carry out their professions (Jeong, 2013). Since teacher candidates' knowledge, abilities and philosophies of LTA practices take form in teacher education programs mainly, it plays an essential role in the development of LAL. Similarly, on the grounds that assessment is an indispensable part of language teaching, competency and professionalism of language teachers, it should be

carried on in-service training programs through workshops, seminars and formal training courses on LTA (Lam, 2015; Volante & Fazio, 2007; Yastıbaş, 2018).

EFL Teacher Education in the World

Good educational practices necessitate qualified teacher education and training, which makes the study of teacher training and INSET significant. According to Burns and Richards (2009), second language teacher education began in the 1960s. In the beginning, courses lasted for short periods, and they were about specific language teaching methods like Audiolingualism. Then, master's programs were offered, and the courses included topics related to theories of learning and methodology as well as practicum. At present, there are different models of language teacher education worldwide. To illustrate, teacher candidates go to colleges to have BA degrees in the relevant major and follow it up with graduate degrees. On the other hand, instead of having a BA degree in language teacher education field, they may be language teachers after getting certain certificates such as CELTA and fulfilling the necessary conditions (Burns & Richards, 2009). However, in some contexts, as Barduhn and Johnson (2005) state, passing certain English proficiency exams is regarded as adequate to be language teachers. As for the other contexts, native speakers of this language could become language teachers regardless of having or not having any experience or certificate.

EFL Teacher Education in Turkey

Teacher education in Turkey has a long history and English Language Teacher Education has been one of most crucial areas in teacher education policies in Turkey. Following recent modifications and adjustments in policies concerning teaching and education of foreign languages, there has been a crucial focus on teaching English intensively to both young learners and adults, which entails the need for qualified and competent English language teachers. To this end, English Language Teacher Education (ELTE) has gone through various changes (Mahalingappa & Polat, 2013) and the Higher

Education Council issued an update in 2017, which describes the structure of initial teacher training program. In the following sections, ELTE training in Turkey will be scrutinized including both pre-service and in-service education.

Pre-Service ELTE in Turkey

Pre-service English language teachers get training at the education faculties of the universities in Turkey and all ELTE programs in all universities follow the curriculum designed by the Council of Higher Education in Turkey (CoHE). While pursuing a standardized curriculum, set by the CoHE, universities can differ from one another in terms of determining elective courses, coursebooks and assessment methods. Council of Higher Education (CoHE) is also in charge of planning, organizing, supervising all higher education institutions in addition to making reforms in the program of different departments incorporating the education faculties in Turkey (Coskun & Daloglu, 2010).

In accordance with the social and global developments as well as educational needs, ELTE programs have also been revised (Grossman, Onkol, & Sands, 2007) and some major changes have taken place in teacher education and the rationale behind these is the quality of education would be increased through eliminating the deficiencies in the existing system with regard to the aims and keeping up with global developments and changes (Kırkgöz, 2017). These changes were mostly related to the content changes instead of systematical ones and the last change took place at the beginning of 2018-2019 academic year. Based on the latest reform, the names and credits of certain courses offered in the ELTE programs have been changed as set by the CoHE. The quantity of pedagogical courses that are based on theories has increased and teaching practice has been placed in the first term of the last year, too. The latest structure provided by the CoHE is shown in the table 1 below:

Table 1*The Structure of ELTE Programs in Turkey*

Content	%
(FK) Field Knowledge	48
(PK) Pedagogical Knowledge	34
(GK) General Knowledge	18

Pre-service ELTE program consists of a four-year training and all ELTE departments employ a standardized curriculum, and three different distincts of courses are provided. They are field knowledge courses (e.g., Teaching English to Young Learners, Materials Adaptation and Development, English Literature I, Linguistics I and II, Language Acquisition, ELT Methodology, Testing and Evaluation in ELT), pedagogical knowledge courses (e.g., Turkish Education System & School Management, Instructional Technology, Educational Psychology, Research Methods in Education, Turkish Education History, Teaching Principles and Methods), and general knowledge courses (e.g., Community Service, Information Technologies, Non-Departmental Electives). Through these courses, ELTE programs seek to prepare prospective teachers as competent and qualified English language teachers that can work at various levels of education including primary and secondary schools and tertiary education.

To be admitted to these programs, students take an exam that is required for admission to a university, and they are placed into universities and departments depending on their exam scores. For ELTE programs, students have to take the English language proficiency exam that includes 80 questions designed in the form of multiple choice. Despite the fact that this exam seeks to assess and evaluate students' language knowledge and skills, it contains questions related to language knowledge including grammar and vocabulary and reading comprehension. In the first year of ELTE programs, language skills of preservice ELT teachers are intended to be enhanced concentrating on each language

skill individually. More specifically, teacher candidates' proficiency in all language skills, which are listening, speaking, reading and writing in English, are targeted to be improved. Besides these, candidates' awareness and competence of pedagogical knowledge are aspired to be developed through courses such as Educational Psychology, Sociology and Philosophy. As for the second year of the program, it can be perceived as a transition period where teacher candidates receive courses such as English literature, Linguistics, Approaches to ELT and Language Acquisition. These courses are intended for candidates to start obtaining knowledge in field courses.

With regard to the third year of the program, it includes both field courses like Teaching English to Young Learners and Teaching Language Skills and pedagogical courses like Turkish Education System & School Management and Measurement and Evaluation in Education. The new curriculum requires candidates to take two courses in Teaching English to Young Learners and Teaching Language Skills, which is different from earlier curricula in that the number of these courses have increased. Besides, unlike the previous programs, Measurement and Evaluation in Education course is started to be offered in that it will contribute to teacher candidates' assessment literacy level overall, which will also promote their language assessment literacy. More specifically, as prospective teachers will learn important concepts of educational assessment including principles of assessment like validity, reliability and practicality as well as scoring and interpreting test results, it will be beneficial for them during the Testing and Evaluation in ELT course and their future professions.

As to the last year, teaching practice is offered in both terms where preservice teachers visit schools and observe the real classroom contexts as well as making demo lessons under the guidance of mentor teachers at schools and supervisors at their department. This course is the sole chance for them to put theory into practice before starting their professional lives. Also, other important courses like Materials Adaptation and Development and Testing and Evaluation in ELT are provided in the seventh and eighth

semesters, respectively. Prior to starting teaching, it is an important competence for teachers to adapt the existing lecture materials taking the needs and levels of learners into account and to create and develop new ones in accordance with their teaching context following the rules they learnt in the course. Moreover, through Testing and Evaluation in ELT course, it is aimed to equip preservice teachers with the sound basis of different assessment methods for different language skills and knowledge as well as designing different types of tests and assessments for different age groups and proficiency levels. The course also includes the principles of assessment besides evaluating test results and giving feedback to learners. In this regard, the course seeks to prepare teachers as competent professionals who can conduct testing and assessment practices and procedures in the classroom accurately.

All in all, the program consists of 48% specialized knowledge, 34% teaching-related knowledge, and 18% general knowledge (CoHE, 2018).

In-Service ELTE in Turkey

In-service training (INSET), which has been a generic term for teachers' professional development, is regarded as one of the most important parts of continuing professional development (CPD). It has been defined as any type of activities in which teachers are engaged to develop their knowledge, skills and professionalism processes following their undergraduate training and to enhance the quality of learning and teaching (Ryan, 1987). In other words, INSET seeks to equip teachers with necessary knowledge and skills required in the 21st century and to fill in the gaps from undergraduate programs. To this end, a variety of activities are held including courses, workshops, seminars, postgraduate programs, conferences and certificate programs. Considering the realities of the globalized world and the need for competency of brand-new skills on the part of teachers as well as pedagogical paradigm shifts, it is extremely important for teachers to follow the developments in many fields such as pedagogical knowledge, methodology and instructional technologies. In this regard, planning and applying a well-designed INSET

program is unquestionably important for conducting effective educational practices, promoting teacher professional development and increasing the learning opportunities.

The education system in Turkey is centralized and the Ministry of National Education (MoNE) implements all educational issues. MoNE organizes and carries out yearly professional development programs for teachers. The General Directorate of Teacher Training and Development is responsible for carrying out INSET and The Department of Supporting and Monitoring Professional Development is in charge of teacher training (Kahraman Özkurt, 2019). All teachers are required to attend INSET programs.

According to the MoNE, two types of INSET programs are available for teachers which are INSET for beginning teachers and INSET for teachers being in the profession for years (MoNE, 1995). While teachers who start teaching attend INSET with the purpose of orientating and adapting, others who have been in the teaching profession for years are given training for enhancing their knowledge and abilities for high quality teaching. Since 2018, the MoNE has highlighted the importance of promoting all teachers to improve their teaching practices and competency in parallel with lifelong learning concept. The range of the activities for teachers include training for professional and personal development and different activities like symposiums and panels arranged with the cooperation of higher education institutions, training for newly appointed teachers, training for the specialist trainers and training for high positions (MoNE, 2018).

It has been announced that INSET activities will be conducted online via EBA (Education and Informatics Network) owing to COVID-19 threat (Milli Eğitim Bakanlığı, 2023b). As for the content of these trainings, they will mostly be about informatics such as Internet of Things Education and Fundamentals of Digital Entrepreneurship.

As for in-service EFL teacher training in Turkey, INSET plans for 2001-2023 have been published by the MoNE (Milli Eğitim Bakanlığı, 2023a). When trainings planned and held for 2001-2023 by the MoNE are investigated, it is found out that there is a variety of the in-service training activities provided to EFL teachers. The examination of them has

revealed that courses and seminars are related to English language teaching methodology and techniques, evaluation and design of coursebooks, curriculum, technology enhanced teacher education and assessment and testing techniques. The information provided in Table 2 summarizes the accessible in-service activities for EFL teachers according to the years they are offered, the subject matter they incorporate and the intended participants.

Table 2

List of In-service Trainings that are available to EFL Teachers in Turkey

Year	Subject Matter	Participants
2001	Seminar on ELT Methodology	EFL Teachers working in Bitlis, Bingöl, Bolu, Gümüşhane, Hakkari, Kayseri, Kocaeli, Muş, Siirt, and Şırnak city centers; English Teacher Trainers; EFL Teachers working at Anatolian Teacher Training High Schools (15 years of experience); EFL Teachers working at country schools in Eastern and Southeastern Anatolia
	EFL Teacher Trainer Course	EFL Teachers scoring A and B level in the language test (KPDS)
	EFL Teacher Training Project Preparatory Course	EFL Teachers
	Seminar on European Language Portfolio	EFL Teachers
2002	Seminar on ELT Methodology	EFL Teacher Trainers
	EFL Teacher Trainer Seminar (development)	EFL Teachers scoring A and B level in the language test (KPDS) and 5-15 years of experience
	EFL Teacher Trainer Course (training)	EFL Teachers working in Bursa, Çorum, Eskişehir, Sakarya, Samsun and Trabzon city centers as group leaders; EFL Teachers working in Ağrı, Adıyaman, Amasya, Iğdır, Kırşehir and Tokat city centers; Teachers giving English lesson in Adıyaman city center; EFL Teachers working at Science High Schools
2003	Seminar on ELT Methodology	EFL Teachers working in Aydın, Kahramanmaraş, Kırıkkale, Kütahya, Mersin and Sivas city centers as group leaders; EFL Teachers working in Bilecik, Burdur, Elazığ, Kilis, Niğde, Osmaniye and Yozgat city centers
	Seminar on Foreign Language Activities	EFL Teachers
	EFL Teacher Trainer Seminar (development)	EFL Teacher Trainers
	Seminar on Preparing English Teaching Documents	EFL Teacher Trainers
	Seminar on Preparing English Coursebooks	EFL Teachers
2004	Seminar on ELT Methodology	EFL Teachers working in Ankara, Antalya, Ardahan, Artvin, Balıkesir, Çanakkale, Denizli, Düzce, Gaziantep, İzmir, Karabük, Karaman, Manisa, Mardin, Osmaniye, Uşak and Yalova city centers

Table 2 (cont'd)

Year	Subject Matter	Participants
2005	Seminar on EFL Methodology	EFL Teachers working in Adana, Afyon, Batman, Bartın, Bolu, Çankırı, Diyarbakır, Kars, Kastamonu, Konya, Ordu, Şanlıurfa, Rize and Zonguldak city centers; EFL Teachers working in Rize and Yalova county schools (3-15 years of working experience)
	Seminar on Evaluating English Coursebooks	Teachers responsible for preparing English coursebooks in İstanbul and İzmir
	Seminar on Introducing Prep-class English Coursebooks	2 English coordinator teachers from each province
	English Course	Newly appointed EFL Teachers working in Science and Social Sciences High Schools (the ones who did not take this course)
	EFL Teacher Trainer Seminar (Level 2)	EFL Teachers participating in Level 1 of the course in 2004
2006	Seminar on EFL Methodology	EFL Teachers working in Aksaray, Bayburt, Erzincan, Giresun, Gümüşhane, Hatay, Kayseri, Kırklareli, Malatya, Muğla, Nevşehir, Sinop and Tekirdağ city centers; EFL Teachers working in Rize county schools (2-5 years of working experience)
	Seminar on Preparing and Evaluating English Coursebooks	EFL Teachers who are responsible for preparing English coursebooks
	Project Preparing Techniques Course	EFL Teachers working at Technical Education for Girls High School who were suggested to work on a project by the school administration
	Seminar on Introducing Prep-class English Coursebooks	EFL Teachers working at Anatolian, Anatolian Fine Arts, Science and Social Sciences High schools (2 Teachers from metropolises and 1 Teacher from other cities)
	EFL Teacher Trainer Seminar	English Teacher Trainers
	Seminar on Evaluating the Foreign Language Lesson Curriculum	EFL Teachers (2 Teachers from metropolises and 1 Teacher from other cities)
	Course on Techniques and Methods in English Language Teaching	EFL Teacher Trainers
	DYNED English Language Teaching Seminar	EFL Teachers working in primary schools (Ankara and İzmir)

Year	Subject	Participants
2007	Seminar on ELT Methodology	EFL teachers working at primary and secondary schools in Bursa, Edirne, Erzurum, Eskişehir, Isparta, and Kocaeli city centers; EFL teachers working at secondary schools in Amasya and Trabzon; EFL teachers working at primary schools (4 th and 5 th year) in Erzurum (2-5 years of experience)
	Project Preparing Techniques Course	EFL teachers working at Technical Education for Girls High Schools who were suggested to work on a project by the school administration
	Teaching Methods and Techniques Course	EFL teachers working at Anatolian, Anatolian Fine Arts, Science and Social Sciences High schools
	Seminar on Introducing DYNED English Language Teaching System	EFL teachers (Aksaray)
	Practical English Speaking Methods and Techniques Teacher Trainer Training Course	EFL teacher trainers
	Seminar on DYNED English Software	EFL teachers working at primary schools in Gaziantep
2008	Seminar on ELT Methodology	EFL teachers working at primary schools in Istanbul (Anatolian Side); EFL teachers working at secondary schools in Burdur, Çorum, Kırşehir, Samsun and Tokat city centers; EFL teachers working at secondary schools in countries (2-5 years of experience)
	Seminar on Curriculum Introduction	EFL teachers working at piloting schools within the scope of Secondary Education Project
	Seminar on Introducing DYNED English Language Teaching System (Primary Schools) (Cancelled)	EFL teachers who are representatives of DYNED provincial district (Erzurum, Mersin, Rize); EFL teachers working at primary schools; EFL teachers working at primary schools with MA and PhD degrees in English or primary school teaching
	Seminar on English Curriculum Introduction (Cancelled)	EFL teachers working at primary schools, EFL teachers working at primary schools with MA and PhD degrees in English or primary school teaching
	Seminar on Primary Level English Curriculum Introduction and Teaching Methods	EFL teacher trainers
	DYNED English Language Teaching Seminar	EFL teachers working at primary schools
	EFL Teacher Trainer Course	EFL teachers working at Anatolian Hotel Management and Tourism Vocational High Schools
	English Course	EFL teachers working in Ankara city center
	Online Seminar on ELT Methodology	EFL teachers working at primary schools in Istanbul (European Side), Ankara city center
	Seminar on English Language Teaching Techniques and Assessment and Evaluation	EFL teachers working in Ankara In-service Training Institute
	Seminar on Preparing Professional English Teacher's Guidebook	EFL teachers

Table 2 (cont'd)

Year	Subject	Participants
2009	Seminar on ELT Methodology	EFL teachers working at secondary schools in Balıkesir, Bilecik, Çanakkale, Kahramanmaraş, Karaman, and Mersin city centers; EFL teachers (2-5 years of experience); EFL teachers working at Commerce and Tourism Education Schools)
	Seminar on DYNED English Use and Applications	EFL teachers working at primary schools
	Seminar on Curriculum Introduction	EFL teachers working at Technical Education for Girls High Schools, Technical Education for Boys High Schools, and at schools under General Directorate of Commerce and Tourism Education, and General Directorate of Health Affairs and Teacher Education
	Seminar on Introducing English Coursebooks Assessment and Evaluation Seminar	EFL teachers working at Anatolian High Schools
	Seminar on Classroom English	EFL teachers working in 300 General and Vocational Education High Schools that will be subject to final application; EFL teachers; EFL teachers working at General High Schools
	Seminar on English Curriculum Introduction	EFL teachers working at secondary schools
	Online EFL Teacher Trainer Course	EFL teachers working at secondary schools in countries (2-5 years of experience)
	Seminar on Language and Cultural Differences Education	EFL teachers working at primary schools in Aksaray, Bolu, Çankırı, Eskişehir, Karaman, Kırşehir, Kırıkkale, Konya, Nevşehir, and Niğde
2010	Web-based Content Development Course (Level 1)	EFL teachers working at primary schools in Ankara city center
	Web-based Content Development Course (Level 2)	EFL teachers working at primary schools who can use IT tools well
	Seminar on Evaluating Student Achievement Determination Exam 2008 (ÖBBS-2008) Report	EFL teachers working at primary schools (the ones who participated in Level 1 of the course)
	English Curriculum, Teaching Methods and Techniques Course (Level 1)	EFL teachers working at primary schools where the exam was administered
	English Curriculum, Teaching Methods and Techniques Course (Level 2)	EFL teachers working at primary and secondary schools
	Seminar on DYNED English Language teaching System	EFL teachers who participated in and successfully completed the Level 1 of the course
	Seminar on Curriculum Introduction	EFL teachers working at primary schools
		EFL teachers, who are group leaders working at Anatolian and Social Sciences Schools with preparatory schools under the General Directorate of Secondary Education

Table 2 (cont'd)

Year	Subject	Participants
2011	Seminar on Sharing Student Achievement Determination Exam 2008 (ÖBBS-2008) Report	EFL teachers
	Seminar on DYNED English Language Teaching System	EFL teachers selected ex officio by province coordinators, Province coordinators (selected ex officio)
	Seminar on Evaluating EFL Teacher Training	EFL teachers participated in EFL teacher trainer course in 2010 and selected ex officio
	Seminar on Sharing Student Achievement Determination Exam 2009 (ÖBBS-2009) Report	EFL teachers working at schools that were part of the project
	Seminar on Material Design	EFL teachers working at secondary schools
	Seminar on Assessment and Evaluation Techniques	EFL teachers working at secondary schools
	Seminar on New Approaches to ELT	One EFL teacher selected by each of the Anatolian Teacher Training High School Leader Teachers Board
	Seminar on Project Preparing Techniques	EFL teachers working at Science and Art Centers who are selected ex officio to prepare projects about gifted students
	Vocational English Teaching Methodology Course	EFL teachers working at Commerce and Tourism Education Schools
2012	Seminar on DYNED English Teaching System	EFL teachers
2013	Seminar on Common European Framework of Reference for Languages (Cancelled)	EFL teachers who are licensed to administer TELC European language certificate examinations; EFL teachers working at institutions under General Directorate of Lifelong Learning
	Seminar on Project Preparing Techniques (Cancelled)	EFL teachers working at institutions under General Directorate of Lifelong Learning
2014	Seminar on ELT Methodology and Techniques	EFL teachers selected ex officio amongst the ones working at schools under the General Directorate Vocational and Technical Education
2015	None	None

Year	Subject	Participants
2016	Seminar on Updated English Curriculum (Cancelled)	EFL teachers working at schools under General Directorate of Primary Education
	FATİH Project- Technology Enhanced EFL Teacher Trainer Course	EFL teachers who participated in FATİH project Technology Use Course
	FATİH Project- Lesson Procedure Design Course	EFL teachers working at schools under the General Directorate of Secondary Education and General Directorate of Primary Education, especially the ones who participated in Technology Use in Technology, Teaching Methods and Techniques in-service training courses
	Seminar on ELT Methodology and Techniques	EFL teachers working at Anatolian Theology (Imam-Hatip) and Theology secondary schools, the ones who did not participate in in-service training courses offered by the General Directorate of Religion Education within the last three years (2013 and later)
	Updating English Language Teaching Course	EFL teachers selected ex officio
2017	FATİH Project- Technology Enhanced EFL Teacher Trainer Course	EFL teachers who participated in FATİH Project Interactive Classroom Management Training
	Seminar on DYNED English Language Teaching System	EFL teachers working at middle schools under the General Directorate of Primary Education
2018	FATİH Project- Field Specific Lesson Procedure Design Course for Teacher Trainers	EFL teachers participated in FATİH Project Interactive Classroom Management Training who work at primary schools and secondary schools
	New Approaches to Foreign Language Teaching Teacher Trainer Education (Cancelled)	EFL teachers selected ex officio by the General Directorate of Teacher Training and Development amongst the ones working at MoNE schools
2019	Science and Production Oriented Learning Teacher Trainer Education Course	EFL teachers selected ex officio amongst the ones working at MoNE schools in Hatay, Sivas, Uşak, Van and Zonguldak provinces.

Table 2 (cont'd)

Year	Subject	Participants
2020	None	None
2021	Seminar on Methods and Techniques for Effective English Teaching for Students with Different Abilities	EFL teachers working at schools under General Directorate of Teacher Training and Development
	Seminar on Methods and Techniques for Online Assessment and Evaluation of English Speaking Skills	
	Seminar on Self-Assessment Methods in English Language Teaching	
	Seminar on Project-Based Learning in English Language Teaching	
	Seminar on Assessment and Evaluation in English Language Teaching; Learning Styles and Positive Feedback	
	Seminar on Digital Literacy for English Teachers	
2022	Seminar on Assessment and Evaluation in English Language Teaching; Learning Styles and Positive Feedback	EFL teachers working at schools under General Directorate of Teacher Training and Development
	Seminar on Innovative Thinking in English Language Teaching	
	Seminar on Developing English Language System Instruction Skills	
	Professional Development Communities (PDC) Educator Training Course for English Teachers	
2023	Seminar on Assessment and Evaluation in English Language Teaching; Learning Styles and Positive Feedback	EFL teachers working at schools under General Directorate of Teacher Training and Development
	Seminar on Innovative Thinking in English Language Teaching	
	Professional Development Communities (PDC) Educator Training Course for English Teachers	
	Seminar on Developing English Language System Instruction Skills	

As stated by Mirici and Pulatsü (2022), it is recognized that undergraduate education alone does not offer a sufficient knowledge foundation for the teaching profession, emphasizing the substantial importance of INSET programs in enhancing teachers' knowledge and skills. In light of this, it is necessary to regularly review, and update INSET programs to establish a strong educational foundation for teachers that encompasses various aspects, including Language Testing and Assessment (LTA). These programs should be designed to effectively address the specific needs of English as a Foreign Language (EFL) teachers in the field of LTA. On the other hand, INSET in Turkey has been addressed as inadequate for catering to the needs of language teachers in the literature

(Günel & Tanrıverdi, 2014; Küçüksüleymanoğlu, 2006; Özer, 2004; Uysal, 2012; Uztosun, 2018). The problematic aspects involve both the number and the quality of training programs including the plan, application and review of trainings. As stated by Küçüksüleymanoğlu (2006), in-service training courses given to English teachers from 1998 to 2005 were not adequate with regard to the number of INSET courses for all teachers irrespective of their majors. In a similar vein, the study of Uztosun (2018) investigated the beliefs of English teachers on the strong and weak sides of in-service training programs in which they participated. It was found that the number of the available programs was few and there were some problems in different stages of the courses offered such as the plan, implementation, follow-up and evaluation of the courses.

Especially, planning of the courses is stated to be problematic (Özer, 2004; Uysal, 2012; Uztosun, 2018). More specifically, needs analysis has not been carried out prior to these programs, which causes the scope of the programs not to be related to the actual needs of teachers. Therefore, a top-down approach has been adopted in terms of the determination of the scope of the course without analyzing the actual needs of teachers. In this regard, it is of paramount importance to state that planning of these programs should focus on bridging the gaps and including relevant needs of teachers.

Another concern is related to the implementation of the courses in that the most frequently used instructional method is lecture instead of others such as discussions and reflective practices. Traditional lecturing is not stated to be well suited to the needs of teachers attending INSET programs since there is a greater focus on theory rather than practice (Bayrakçı, 2009; Uztosun, 2018). Besides, teachers articulated that the instructors giving these INSET courses were not prepared well and adequate, which is an important shortcoming. Teachers also stated that timing of these programs was not appropriate for their timelines and regarded the location of the programs as inconvenient for practice purposes. Moreover, the review of the literature points to the lack of follow-up and evaluation of the trainings (Uysal, 2012) and teachers highlighted the importance of them so that they

could reflect on important issues and talk to course instructors about different issues like how to put theory into practice, which will contribute to the effectiveness of such training programs.

Studies on LAL of in-service and pre-service EFL Teachers in Turkey

The review of literature has indicated that very few studies have investigated the language assessment literacy (LAL) of in-service and pre-service EFL teachers in Turkey in spite of the fundamental role of LAL on language education. There are a limited number of studies on the training levels and needs in language testing and assessment of in-service EFL teachers as well as their beliefs and practices of assessment (Cirit, 2015; Han & Kaya, 2014; Haznedar, 2012; Karagül, Yüksel, & Altay, 2017; Köksal, 2004; Mede & Atay, 2017; Ölmezer-Öztürk & Aydın, 2018a, 2018b; Yastıbaş, 2018; Yastıbaş & Takkaç, 2018) and training levels and needs of pre-service EFL teachers with regard to LAL and the evaluation of assessment and testing course as well as the preservice teachers' beliefs towards this course (Hatipoğlu, 2010, 2015; Tamerer, 2019).

Hatipoğlu (2010) investigated the English Language Testing and Evaluation course offered in the last year of language teacher education programs in Turkey besides the opinions of prospective EFL teachers towards this course. More specifically, this study aimed to explore the English Language Testing and Evaluation course (ELTEC) by focusing on the scope and methodology of the course, which provides a summative evaluation of this course. To this end, a questionnaire was used, and interviews were carried out with prospective EFL teachers from a state university in Turkey. In the questionnaire, 81 participants were asked to state five subjects that were included in the course and the benefits of these subjects for their future professions as well as providing the reasons. Also, they were asked to specify three issues that should be altered related to the course through listing the reasons. As for the interviews, they were carried out with 16 prospective EFL teachers to get more specific information. The study showed that assessment of language

skills and areas and principles of testing such as validity and reliability were regarded as the most beneficial topics for their future careers. In addition to these, preparing test items like multiple choice items and types of testing were indicated. Also, prospective EFL teachers regarded some subjects provided in the course as not concrete and practical. Therefore, they highlighted the significance and necessity of giving a place to practical components in language testing and assessment, which will contribute to the enhancement of LAL. Besides, the coursebook utilized in the course and other supplementary readings were stated to be too complex for them to grasp and follow.

In 2015, Hatipoğlu carried out another study with the aim of revealing the needs and opinions of prospective EFL teachers in Turkey with regard to language testing and assessment. In other words, 124 prospective EFL teachers' LAL levels and expectations from the course were sought to find out. To this end, needs analysis survey questionnaire and interviews were used to EFL teacher candidates. It was found that they took only one testing and assessment course during their education and apart from two candidates all teacher candidates expressed the essential role of this course for them. Despite this, almost half of them regarded only one assessment course as sufficient and contribute to their development in assessment field. On the other hand, many participants did not offer any subjects to be included in the future courses. As for the others, the minority of them suggested some subjects to be added to the courses, which are alternative assessment like journals, portfolios and self-assessment, administration of exams like time allocated and using statistics for evaluating. All in all, it was suggested that the significance attached to language testing and assessment courses and training in this field should be increased and teacher educators should collaborate with teacher candidates for enhancing the quality and effectiveness of such courses.

In another study, Mede and Atay (2017) examined LAL of EFL instructors who were working at preparatory schools of universities in Turkey. The study employed both quantitative and qualitative research methods, with the quantitative aspect involving the use

of an online questionnaire and the qualitative aspect involving focus group interviews. EFL teachers were found to have limited LAL levels and to need further training in three fields of language testing and assessment. Of all areas, classroom focused language testing and assessment was the most stated area for further training and as for the purposes of testing, the participants regarded themselves as sufficient in this area. They also specified a further training need in assessment of speaking skill. Moreover, participants indicated that the training they received was not adequate since the training was focused mostly on exam and for once. In light of these findings, it was suggested that in-service training in language testing and assessment should incorporate an intense focus on putting several forms of assessment into practice in preparatory schools. Moreover, the needs of EFL teachers working at preparatory schools should be considered to increase LAL level through cooperative trainings.

On the other hand, Tamerer (2019) explored the prospective EFL teachers' levels of training and further needs in LAL in Turkey. A questionnaire was administered to 30 teacher candidates and interviews were conducted with 10 participants. It was found out that these participants had low levels of LAL and they specified a basic training in fields of language testing and assessment. It was suggested that the findings are important for increasing the awareness related to the significance of assessment and testing in language education. Besides, some changes can take place regarding assessment courses by taking the needs of teacher candidates into consideration. More specifically, certain activities like micro-teachings, conferences and workshops are offered for the inclusion in courses.

In his doctoral dissertation, Yastıbaş (2018) studied the way language assessment is applied in language classroom settings in Turkey. Through interviews, observations, focus group discussion, think-aloud protocol and document analysis, the author collected data from 8 instructors teaching EFL at a university. A multiple-case study research design was employed in this research. The results of the study showed that these instructors regarded assessment and evaluation in a critical manner. Also, it was revealed that those

who got their undergraduate training in education faculties regarded their assessment training as efficient; however, other instructors who graduated from different faculties mentioned inefficient assessment training in university. Furthermore, these instructors stated that they enhanced their language assessment and evaluation knowledge building upon their experience. More specifically, it was pronounced that once before their professional life, they were not very knowledgeable about the area of language assessment. To compensate for their weaknesses, they either employed the ways they were assessed as learners or included peer assessment and collaborated with their co-workers. Lastly, the results demonstrated that these participants improved their own descriptions of some basic terms regarding assessment such as reliability and validity. With regard to validity, they generally took content validity into consideration and their perception of reliability was connected with reliability. In relation to these findings, it was suggested that there was a need for professional development program on LAL, which should encompass the essential components of LAL. These constituents ought to be clarified in a comprehensible manner and promoted by authentic instances. As for the statistics in assessment, it was not suggested to be included in the program on the grounds that the instructors did not use any statistical analysis in their classrooms. Another recommendation made in the study is that a textbook regarding language assessment field can be generated for individual work and preservice training lectures, and the book should cover both theoretical and practical aspects of assessment.

Studies on LAL of in-service and pre-service EFL Teachers Abroad

Researchers have sought to investigate LAL in the past two decades since they grasped the cruciality of language testing and assessment (Fulcher, 2012; Jeong, 2013). The review of literature has demonstrated that there has been an increasing number of studies on LAL focusing on in-service and pre-service EFL teachers' LAL levels and their expectations regarding further training in LTA (DeLuca & Klinger, 2010; Hasselgreen,

Carlsen, & Helness, 2004; Herrera & Macías, 2015; Inbar-Lourie, 2008; Jin, 2010; Lam, 2015; Vogt & Tsagari, 2014; Volante & Fazio, 2007; Xu & Brown, 2017).

To start with Hasselgreen, Carlsen, and Helness' (2004) seminal study, they sought to find out training received in LTA and training needs of three various respondents who are teacher educators, language teachers and assessment specialists from different countries. In line with that purpose, an online questionnaire was employed to 914 respondents and 361 of them were teachers. Three areas of LTA that were classroom-focused assessment, contents and concepts of LTA and purposes of testing were included in the questionnaire. The results demonstrated that training in LTA was problematic and not sufficient, and there was a crucial need for further training in LTA, which was stated by all three respondent groups. To start with classroom-focused assessment area, many language teachers asserted that they did not get organized training in many aspects like giving feedback, adapting prepared tests and informal assessment. In a similar vein, they stated that they did undergo formal training in the purposes of testing area; therefore, they did not feel confident about certain subjects like placement of students. As to the last area, which is contents and concepts of LTA, the majority of teachers lack training in many fields such as using statistics and ensuring reliability. Therefore, the need for training was expressed explicitly in all three areas mainly in two areas: concepts and contents of LTA and purposes of testing. However, it was of the utmost importance to consider certain points in the interpretation of the results since language teachers were not only teaching in the classroom but also applying tests and various forms of assessment as experts or teacher educators. Also, it was argued that these teachers could have some background in language testing and assessment, which could cause misinterpretation of the target participants. In light of this information, it was proposed that exploring language teachers who work only as classroom teachers might be more useful for improving the validity of findings and getting accurate data for the LAL of language teachers.

Vogt and Tsagari's (2014) study that this PhD dissertation mainly draws on sought to investigate the LAL of language teachers and their training needs in the LTA field. Having adapted the questionnaire of Hasselgreen et al.'s (2004) questionnaire, they employed this to 853 EFL teachers from various European regions including Turkey. Following the questionnaire, they had an interview with 63 EFL teachers. More specifically, through questionnaires it was aimed to uncover the training they received and needed in various domains of LTA and via interviews it was sought to find out their individual needs in these domains and to specify the ways they employed to make up for problems in specialty in these areas. The findings demonstrated that many of the respondents did not receive advanced training in three domains of LTA and they highlighted the need for further training in these three domains. In this regard, the findings of this study were in line with those of Hasselgreen et al. (2004)' study. They also stated that they received the least training in the domain of purposes of LTA and to make up for this inadequate training, they tried to learn it in their professions. At that point, the researchers signified that though EFL teachers could handle this problem partially, they might restrain themselves from learning current ways of language assessment; therefore, they recommended that in-service trainings should involve such up-to-date methods and practices of assessment in their curricula. Moreover, it was found that the majority of EFL teachers stated that their training in LTA did not suit the realities of classroom context. As their training mainly included conventional forms of assessment and the classroom context required them to carry out diverse forms of assessment, they underscored the insufficiency of their training in the field of assessment in teacher education programs. Besides, alternatives in assessment were indicated as one of the areas that they lack adequate knowledge and practice and stated as one the crucial topics to be included in in-service training programs.

In 2010, DeLuca and Klinger explored the opinions of preservice teachers with regard to assessment training they received. A questionnaire was used to investigate preservice teachers' confidence in terms of LTA theories, practices and philosophy in

addition to what they expected from the assessment course offered in teacher education programs. In total, there were 288 teacher candidates and the analysis of the findings revealed that taking assessment course has increased their confidence levels. More specifically, it was found that confidence level in theories and practices of assessment was higher compared to that of in philosophy of assessment. Another important finding of the study was that some subjects like reliability and validity of items and revealing the philosophy of their assessment practices as well as adapting certain assessment methods to their contexts were offered to be incorporated into that course, which would contribute to assessment literacy of teacher candidates.

In a different study, Lam (2015) aimed to investigate training in language testing and assessment area in teacher education programs in terms of opinions of both teacher candidates and educators on to what extent assessment courses contribute to or prevent the enhancement of LAL in Hong Kong. For this purpose, both survey and interviews were utilized for collecting data in that survey was used in all teacher education programs and focus group interviews were carried out with 9 teacher educators and 40 teacher candidates. Also, course outlines and program guides were analyzed. The findings showed that training in LTA was not adequate and failed to connect the theory and practice of LTA. Moreover, the training was regarded as too technical and ethics and justice in testing were not involved in this training. As for the practice component, many of teacher candidates stated their concerns about the usefulness of LTA knowledge in their practicum and their future teaching professions on the grounds that mentor teachers in school settings were said not to employ alternatives in assessment, but to use conventional methods of assessment. Besides, it was asserted that LTA knowledge of teacher candidates should be assessed in practicum so that they would not ignore the fundamental role of assessment in education. In short, Lam (2015) highlighted that involvement of LAL in assessment courses is of pivotal importance to enhance LAL of teacher candidates and to equip them with essential knowledge and skills of language assessment.

In a similar study, Volante and Fazio (2007) analyzed the enhancement of LAL of preservice teachers in Canada via questionnaire employed to 69 respondents. They aimed to delve into how pre-service teachers regarded their level of assessment literacy and explore the implementation and purposes of assessment methods and their further training expectations. The findings indicated that their self-perception regarding their level of AL was low and as for assessment purposes they were more inclined to prefer summative assessment than formative assessment. Similar to earlier studies, it was suggested that there was a need for more training in alternatives in assessment like journals and portfolios instead of conventional methods of assessment. Therefore, Volante and Fazio (2007) underscored the importance of conducting needs analysis regarding the gap in assessment knowledge and practices of teacher candidates and language teachers. They concluded that offering courses on assessment does not necessarily mean that teachers were equipped with solid foundations of assessment knowledge and skills. Shortly, the results of Vogt and Tsagari (2014) and Lam (2015) were corroborated.

In 2019, Sultana investigated LAL of English language teachers in Bangladesh and the main aim of the study was to grasp and measure the assessment literacy of language teachers. The study sought to investigate whether they were well-prepared for carrying out tasks related to assessment and how they comprehended LAL in their teaching. Semi-structured interviews were carried out with 10 secondary English teachers who work in five different school setting. Analysis of qualitative data revealed that background of these teachers was not enough for performing assessment practices accurately and the training they received in undergraduate education was not sufficient since only basic terms of testing and assessment were covered during the testing course. Furthermore, they stated that their undergraduate training overwhelmingly covered issues related to learning how to teach rather than assessment and testing. Considering that limited training, knowledge and skills of language assessment were shaped and enhanced during their teaching practices through experience. Based on the findings, it was suggested that language testing and

assessment should be included more in teacher training programs, which would contribute to ongoing professional development of teachers. In addition, a language testing organization was recommended to be founded to offer diverse opportunities in the assessment field both for teachers and test developers.

Along the same lines, Fard and Tabatabaei (2018) delved into the level of LAL of 52 EFL teachers working in both state and private schools in Iran. They sought to find out the extent to which these teachers were well-informed about assessment. For this purpose, "Classroom Assessment Literacy Inventory" was employed to them. According to the findings, their literacy levels were not high enough and should be increased. In addition to these, the participants did not find themselves prepared for dealing with difficulties in classroom assessment and did not have the awareness of what high quality assessment connotes.

In Brazil, Duboc (2009) conducted a study for investigating the perceptions and practices of EFL teachers working in primary school setting. The data were collected through interviews, observations and tests from 3 in-service teachers. The study showed that there were three major difficulties. While the first one was related to the issue that the concepts of measurement and evaluation were comprehended as identical, the second problem was that these instructors disregarded some areas of language development like language skills and instead of these, they focused much on language knowledge, that is to say, vocabulary and grammar. Regarding the last issue, the participants were expressed to have used written tests by a majority instead of other forms of assessment like alternative. Thus, teacher education programs were suggested to be planned once again for the betterment of LAL training to cater to the needs of language teachers.

In another study, Wach (2012) investigated EFL teachers' opinions and classroom practices in assessment in Poland. Whether being a native speaker or nonnative speaker of English and the type of school influence their beliefs of assessment. To this end, 42 university teachers and 45 school teachers participated in the study. Among university

teachers, 12 were native speakers. Out of 45 school teachers, there were primary, middle and secondary school teachers (17, 20, 8, respectively). Data was collected through questionnaire, and it was found out that compared to the school teachers, awareness of university teachers of assessment fields and functions was higher. Another point found is that these participants regarded assessment and instruction as independent of each other. Also, almost all concentrated on testing grammar and vocabulary and being native or nonnative did not change much in terms of the beliefs of them. Consequently, it was recommended that training should be provided to EFL teachers covering a variety of topics like diverse language skills.

By means of investigating the preparedness of 43 ELF teachers who work at different grade levels to assess and the kind of problems they experienced in classrooms in Indonesia, Hudaya (2017) found that more than 50% of them evaluated themselves as ready to conduct assessment to gauge learners. Moreover, the majority of them asserted that the principles of assessment which are practicality, reliability, validity, authenticity and washback were significant. On the other hand, they had some problems in providing feedback to learners. Overall, it was found that these teachers had assessment literacy.

Another recent research on LAL was carried out in China by Xu (2017) through a 3-year longitudinal study with inexperienced EFL teachers. During this time period, LAL of EFL teachers were monitored through observations, interview, journal entries and fieldnotes. The analysis of these data revealed that there were 3 stages of improvement of these beginning teachers. Whereas participants started to learn practical skills for predetermined assessment in the initial stage, their mindfulness of the relationship between goals and assessment applications increased in the middle stage. As for the last stage, their ability to conduct assessment without much preparation progressed particularly in formative assessment. By underlining the importance of LAL for EFL teachers, the researcher concluded that being able to design and carry out assessment spontaneously was the

indicator of LAL development, and these teachers enhanced their assessment applications in the course of time.

In their study, Vogt, Tsagari and Spanoudis (2020) sought to examine how English language teachers perceive assessment, their levels of language assessment literacy (LAL), and their requirements for training. A survey questionnaire was completed by 113 teachers from Germany and 379 teachers from Greece. The data collected were analyzed using various statistical methods such as repeated measures analysis of variance (RM ANOVAs), correlation analyses, and confirmatory factor analysis. Additionally, qualitative data obtained from interviews with 25 German and 20 Greek teachers were used to provide further support. The findings revealed that teachers share similar perspectives on assessment and LAL concepts, but their perceived training needs differed based on their educational contexts. The interviews provided valuable insights into the contextual factors influencing these perceptions. Especially, on average, 60.1% of the participants expressed a need for "a little" or "advanced" training in Language Testing and Assessment (LTA), with specific topics varying based on the local educational settings. In contexts where high-stakes testing is prominent, such as Greece, respondents requested more advanced training, as observed in previous research by Vogt and Tsagari (2014). Comparing our study to the findings of Hasselgren et al. (2004), where participants reported low levels of Language Assessment Literacy (LAL) and expressed a general need for LTA training, teachers in our study similarly exhibited low LAL levels (35% reported no training and 32.5% reported minimal training in LTA). Moreover, teachers in both contexts identified similar areas for training, namely alternative assessment formats like peer assessment, self-assessment, and portfolio assessment. Other aspects requiring development included testing micro-linguistic aspects and language skills, as well as grading using non-traditional assessment methods. Establishing quality criteria for assessments, such as reliability and validity, was found to be lacking. The limited formal assessment training among teachers is a recurring finding supported by various studies conducted in different educational contexts.

Various contextual factors at different levels interact and influence teachers' perceptions of their training needs in Language Assessment Literacy (LAL) and their assessment practices, as evidenced in the questionnaire data and further explored in the interview data. The findings indicate that collaboration among teachers plays a crucial role in compensating for the lack of formal training and in developing a practical skill set related to LAL. The interview data reveals a strong interest in the implementation aspect of perceived training needs. This suggests that LAL training initiatives aimed at meeting the needs of language teachers should take into account their existing knowledge and practices, leveraging and enhancing them to positively impact their assessment methods. The findings indicate that when designing training programs to improve the language assessment literacy levels of teachers, it is essential to consider the diverse contextual factors, characteristics, needs, and traditions. Despite lacking formal assessment training, teachers possess assessment-related experience. Therefore, training programs should make use of and incorporate teachers' existing assessment experiences. Additionally, the significant role of collaboration at the intermediate level in the data suggests that training measures should also incorporate collaborative elements that foster shared reflection on assessment practices across different educational contexts.

In her dissertation, Al-Bahlani (2019) aimed to examine the status quo of LAL of EFL teachers in Oman in addition to the assessment theory and applications in real classroom settings. Mixed-method research design was used in that self-assessment surveys, a language assessment knowledge test, an assessment evaluation task, classroom observations and teacher interviews. To analyze the gathered data, multivariate analyses were employed to investigate the connection between LAL and different variables like experience of teaching, pre- and in-service trainings, gender, and academic background. The results demonstrated that there was both harmony and disharmony between their self-assessed LAL and showed assessment knowledge. As to the perceptions of these teachers regarding their competency in LAL, it was partial. Also, among the variables that were

measured, preservice training in assessment was the most influential variable on their LAL. In this regard, it was pointed out that assessment courses given in preservice education was critical for sound assessment practices and knowledge, and the current programs were suggested to be revised and re-evaluated taking the immediate necessity of elaborating on these programs into account. More specifically, teachers should be trained in many domains of assessment encompassing design of rubrics, giving feedback, providing practical courses on how to create digital assessment tools, formative assessment as well as self- and peer-assessment. It was especially highlighted that the teachers were not able to give feedback to learners over the course of various assessment; therefore, they ought to be made conscious of pros and cons of excluding or giving feedback. As to the other implications, too much teaching blocks certain assessment applications; thus, it was asserted that teaching load of teachers should be reconsidered. Furthermore, deprivation of technology use in LTA was detected despite having technological devices at schools; thus, the use of technology was stated to be encouraged. Another suggestion was related to the finding that those receiving in-service training and not receiving any in-service training in assessment did not differ from each other much. In line with this, it was stated that in-service trainings provided at present ought to be reviewed and changed to fulfill the needs of EFL teachers.

Chapter 3

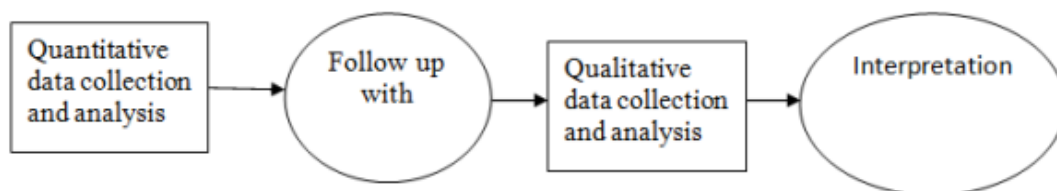
Methodology

Research Design

In this study, the researcher applied a sequential explanatory mixed method design in that the researcher collected and analyzed quantitative data in the first phase and gathered and analyzed the qualitative data in the second phase of the research (Creswell, 2009). In this design, emphasis is placed on quantitative data and qualitative data is used for illustrating and detailing the findings of quantitative data (Lodico, Spaulding, & Voegtler, 2006). In this regard, as Creswell (2012) states, these two different data sets are not collected concurrently in that the analysis of qualitative data is utilized to report and interpret the findings of quantitative data. This mixed method model is advantageous because it has explicitly determined phases, which allows the researcher to carry out each phase respectively. Through this research design, it was aimed to give a general picture of the LAL levels and needs of in-service EFL teachers in terms of LTA and to clarify this picture by giving deeper insights into their needs for further training and their experience in LTA in language classroom context. This mixed method design is shown in Figure 2 below.

Figure 2

Explanatory Sequential Design (Creswell, 2012, p.541)



Mixed methods research refers to gathering, analyzing and integrating both quantitative and qualitative data in order to have a better understanding of research problems through employing quantitative and qualitative approaches (Tashakkori & Teddlie, 2007). Put it differently, combining quantitative and qualitative research and data allows the researcher to answer a great number of research questions and to gain in-depth understanding of the phenomenon under investigation. Also, it enables the researcher to utilize the strengths and cope with the weaknesses of each research method (Johnson & Onwuegbuzie, 2004).

The quantitative phase of the study was a survey study employing a questionnaire as it allows the researcher to reach a vast number of people at a time, which increases the reliability of the study. It is a cross-sectional research in that it will be carried out to collect data from the participants at one time. As for the format of the questions, it was a closed questionnaire consisting of a few open questions. Whereas open-ended questions require the respondents to write an answer, close questions entail them choosing an answer from the given options or grading the statement. In this regard, it is easier to carry out and score the questionnaire since it provides standardized data (Fraenkel & Wallen, 2006). As for the qualitative phase of the study, an interview was adopted since it gives an in-depth understanding of people's perspectives and experience in a specific field (Johnson & Christensen, 2004). Semi-structured interviews were employed in that questions were planned before the interview and the interviewees were able to state their further ideas on questions.

Setting and Participants

The current study was conducted in the formal and non-formal education settings in Turkey. More specifically, it was carried out in different education settings which are public schools, private schools and private language schools. The data for this study were collected from one major informant group: in-service EFL teachers who work at formal and

non-formal education institutions at different grade levels (primary, secondary and high school) in Turkey.

The participants for this study were selected using convenience sampling, which involves choosing individuals who are readily accessible for the research (Mackey & Gass, 2005). As Muijs (2004) states, it is one of the mostly utilized sampling methods in educational research as well as being considerably effective regarding effort, time and money. 300 in-service EFL teachers took part in the quantitative part of the study. Also, 20 in-service EFL teachers participated in the qualitative part of the study.

After eliminating the ones which do not complete all the items in the questionnaire and meet the criteria; in other words, data reduction, there were 300 participants for the first (quantitative) part of the study. Among them, there were 171 female teachers and 129 male teachers. Their ages ranged between 22 to 54 years old. As to their teaching experience, they had various levels of experience as English teachers in that 37.3% of them had 6-10 years of experience, and 31.7% of them had 1-5 years of teaching experience. The rest of them (31%) had 11 and more years of experience in teaching.

Also, 108 of them declared that they are working at private schools while 99 and 93 of them are working at public and private language schools, respectively. They stated to be working at various levels (primary (n=102), secondary (n=101), and high school (n=97)). When they were asked to express their first language, most of them (90.7%) stated it to be Turkish. Other participants listed different languages that are English (4.3%), German (3.3%) and Arabic (1.7%).

Table 3*Profile of the Participants*

	N	%
Gender		
Female	171	57
Male	129	43
Age		
22-32 years	114	38
33-43 years	95	31.7
44-54 years	91	30.3
Education Level		
BA	261	87
MA	27	9
PhD	12	4
Years of Teaching Experience		
1-5 years	95	31.7
6-10 years	112	37.3
11 or more years	93	31
Type of School / Institution		
Public School	99	33
Private School	108	36
Private Language School	93	31
Grade level they teach		
Primary School	102	34
Secondary School	101	33.7
High School	97	32.3
First Language		
Turkish	272	90.7
English	13	4.3
German	10	3.3
Arabic	5	1.7

When it comes to these in-service EFL teachers' taking a testing and evaluation course in their undergraduate education, 81.3% of them declared that they have not taken such a course before. The rest of them (18.7%) stated that they have taken one testing and evaluation course in university education. Also, when these participants were asked to state

whether they have taken any testing and evaluation course in ELT specifically, 97 % of them expressed that they took the course and only 3 % of them stated the lack of that course.

As to their education levels, 261 of them told that they were BA graduates while twenty-seven of them stated that they hold or study for MA degrees and 12 of them hold PhD degrees. The majority of the participants were graduates of ELT departments (n=269) whereas the rest held BA degrees from other departments which are English Linguistics (n=10), English Language and Literature (n=9), Linguistics (n=7), American Culture and Literature (n=4) and Translation and Interpreting (n=1). Including graduates from other departments apart from ELT in the study brings diversity, interdisciplinary collaboration, methodological variations, represents diverse voices, and enriches the implications. Their knowledge and experience can enhance the understanding of language assessment literacy from different angles, enabling a more nuanced analysis of the participants' assessment-related skills, knowledge, and training needs. This approach broadens the scope and depth of the research, fostering a more comprehensive understanding of the subject matter and enhancing the overall quality of the study.

More than half of the participants with MA degrees declared that they studied ELT (n=28). Other specified MA degrees were on Curriculum and Instruction (n=6), English Language and Literature (n=4) and Educational Administration (n=1). With regard to PhD fields, they were ELT (n=8), English Linguistics (n=3) and English Language and Literature (n=1). The details of the BA, MA and PhD graduation can be seen from Table 4 below.

Table 4

Summary of the Participants' BA, MA and PhD Degrees

	N	%	Cumulative %
Undergraduate Major			
ELT	269	89.7	89.7
English Linguistics	10	3.3	93.0
English Language and Literature	9	3	96.0

Linguistics	7	2.3	98.3
American Culture and Literature	4	1.3	99.7
Translation and Interpreting	1	0.3	100.0
MA Major			
ELT	28	71.8	71.8
Curriculum and Instruction	6	15.4	97.4
English Language and Literature	4	10.3	82.1
Educational Administration	1	2.6	100
PhD Major			
ELT	8	66.7	66.7
English Linguistics	3	25	91.7
English Language and Literature	1	8.3	100

As stated earlier in this chapter, after analyzing the data gathered through questionnaire, the participants who were volunteers were invited to take place in the second part of the study and 20 of them agreed to be interviewed. Pseudonyms were given to each interviewee considering ethical issues. The next table indicates the interviewee's background information and the length of interview for every EFL teacher. 6 of the 20 participants have an MA degree. As for their teaching experience, it varies from 1-5 years to 11-15 years. When it comes to the school they are teaching at, 6 of them work at private school and 8 of them work at public school while 6 of them work at private language school. Moreover, their teaching level also differs including primary school (n=10) secondary school (n=6) and high school (n=4).

Table 5*Summary of the interview participants' background information*

Pseudonym	Recording Duration	Education	Years of Experience	Type of School	Teaching Level
Seda	14' 32"	MA	6-10	Private	Secondary
Ahmet	19' 54"	BA	1-5	Public	Primary
Alican	09' 43"	MA	1-5	Private Language	Secondary
Seher	17' 23"	MA	6-10	Private	Primary
Alper	23' 39"	BA	11-15	Private Language	Secondary
Melis	15' 21"	BA	11-15	Private	High
Selin	14' 41"	MA	6-10	Private Language	Primary
Berkin	24' 35"	1-5	1-5	Public	Primary
Doğa	18' 32"	BA	6-10	Public	Secondary
Fatma	13' 49"	BA	6-10	Public	High
Dilan	23' 31"	BA	6-10	Private	Primary
Merve	19' 56"	BA	11-15	Private Language	Primary
Cenk	09' 22"	BA	1-5	Private Language	Primary
Ada	23' 39"	BA	6-10	Public	High
Gizay	13' 48"	MA	1-5	Private Language	High
Mustafa	15' 12"	BA	1-5	Private	Primary
Ayla	18' 29"	BA	6-10	Public	Secondary
Okan	30' 23"	BA	6-10	Private	Primary
Sezen	21' 10"	MA	1-5	Public	Primary
Eda	9' 38"	BA	11-15	Public	Secondary

Instruments

Questionnaire

For the current study, quantitative data were collected using the questionnaire consisting of two parts. The questionnaire was employed to in-service ELT teachers and in the initial part of the questionnaire there were questions like age, gender, education level and years of teaching experience, type of school/institution they teach at and the grade level they are teaching at present. The second part of the were adapted from Vogt and Tsagari's (2014) Teachers' Questionnaire that has three sections and is in the Likert-scale format. These sections are classroom-based language testing and assessment (LTA), purpose of testing, and content and concepts of LTA and each part includes two parts in that Part A refers to the training received, and Part B points out the training needs. In-service EFL teachers were asked to rate the training they think they have received in their undergraduate education in the given items by choosing from a 4-point Likert-type scale (None"; "Little (1-2 days) "; "Sufficient"; "Advanced ") in Part A. In Part B, they were asked to rate the training that they need in the same items given in the previous part by selecting from a 4-point Likert-type scale (None; Basic training; Intermediate training; Advanced training).

This questionnaire was adapted since the items of the instrument are parallel to the information the current study is seeking, which is to explore the current level of in-service teachers' LAL and their training needs in this field. Another reason for adapting this questionnaire is that it has been shown to have high reliability and construct validity proven with high Alpha coefficient (ranging from .80 to .93 for individual scales) following factor analysis process. However, some changes were made in line with the aims of the current study by including a new section "Knowledge of Testing and Assessment" based on Brown & Abeywickrama's (2010) book and excluding some items from three sections as well as making some modifications in the Likert scale to have a better grasp of LAL levels and training needs of in-service ELT teachers (see Appendix-A). After these changes, this

instrument was piloted with a group of participants who are similar to the target group for whom the questionnaire was designed.

Interview

Semi-structured interviews with in-service EFL teachers working in different education settings were conducted and audio taped. The interview questions were adapted from Jeong's (2013) Instructor survey and Instructor interview questions- Language Assessment Course. Some necessary adaptations were made in line with the objectives of the current study by changing the wording of these questions and adding some questions. To validate these questions, they were first reviewed by the research supervisor and two teacher educators' opinions were taken. They were asked to evaluate the wording and clarity of the questions, which provides clarity, content and construct validity as well as redundancy (Brown, 2001). Following this, the instrument was piloted with two in-service EFL teachers to improve their format and control its length.

A semi-structured interview was preferred in this study since it enables the researcher to follow a list of written questions as well as leaving room for necessary changes or additions (Mackey & Gass, 2005). Also, it allows the researcher to ask for more questions when there is a need for clarification and elaboration, which gives the interviewees the chance to state their further opinions. In this regard, it provides more thorough insights and details about their opinions.

The interview form can be seen in Appendix B (Turkish) and Appendix C (English). The interviews were carried out either in English or Turkish based on the interviewee's choices so that they would not feel anxiety during the interviews. There were 11 questions which were open-ended in that the researcher could gather comprehensive data from the participants. The questions were related to their educational background, participants' beliefs of the Testing and Evaluation in ELT course that they received in their undergraduate

education including the sufficiency and scope of the course, implementation of Language Testing and Assessment practices, problems and challenges experienced related to testing and assessment practices as language teachers, possible solutions to deal with these difficulties, needs for training in LTA. The interview questions and their categories can be seen from Table 6 below:

Table 6

Semi-Structured Interview Questions for the In-service EFL Teachers

CATEGORIES	QUESTIONS
1. Educational Background	1) Which departments did you graduate from (BA., MA and PhD)? 2) How long have you been teaching? 3) How many statistics and language testing and assessment course have you taken in your undergraduate training?
2. Opinions and Beliefs about Testing and Evaluation in ELT Course	4) Do you think that Testing and Evaluation Course that you took in undergraduate training covers necessary components of language testing and assessment? 5) Of the topics covered in this course, which topic/s do you think is/are the MOST helpful to you as a classroom teacher? 6) If you were to take this course again, or take an advanced course in language testing and assessment, what topics would you like to learn about?
3. Application of LTA	7) How do you decide which assessment method to use in your classrooms?
4. Needs for training in LTA	8) Do you think that LTA training in undergraduate education is adequate for you as a language teacher? Yes <input type="checkbox"/> No <input type="checkbox"/> If the answer is No, could you state in which areas of LTA do you think you need training? 9) Do you believe it is important for language teachers to be competent at LTA in order to increase the quality of learning and teaching?
5. Problems and solutions to the problems regarding LTA	10) What kind of problems and difficulties you have experienced related to testing and assessment practices as a language teacher (e.g. preparation and administration of tests, students, reliability)? How do you overcome these problems?
6. Additional Comments	11) Do you have any questions or comments about the subject?

All in all, this study employed a questionnaire and a semi-structured interview to gather data and answer the research questions. An overview of the data collection tools, the corresponding research questions, data analysis methods, profile and number of participants can be seen from Table 7 below.

Table 7

Overview of the data collection tools, the corresponding research questions, data analysis methods, the profile and number of participants

Research Questions	Data Collection Instruments	Data Analysis	Statistical Method	Profile and Number of the Participants
1-) How do in-service EFL teachers evaluate their undergraduate training in language testing and assessment? Do they perceive this training as adequate?	Questionnaire	Quantitative	Descriptive statistics	300 In-service EFL teachers
2-) To what extent do in-service EFL teachers perceive a need for in-service training in language testing and assessment?	Questionnaire	Quantitative	Descriptive statistics	300 In-service EFL teachers
3-) Do in-service EFL teachers' training needs in language testing and assessment vary according to a. different educational settings? b. the current level of education that they are teaching? c. year of experience? d. their education level?	Questionnaire	Quantitative	Inferential statistics	300 In-service EFL teachers
4-) How do in-service EFL teachers describe their experience in language testing and assessment practices in classroom context?	Semi-structured interviews	Qualitative	Open and axial coding procedures	20 In-service EFL teachers

Data Collection Procedure

Prior to starting the data collection procedures, permission of the Hacettepe University Ethics Commission was applied for. After getting the approval, in-service EFL teachers teaching at different grade levels and working at various formal and non-formal education institutions were asked to participate in the study through online questionnaire. The researcher gave information about the purpose of the questionnaire regarding the aim of the study and expressed that the data from the questionnaire would not be shared by any other people or institutions as well as highlighting that it would be used only for the purpose

of this study. Also, they were asked to sign the official consent form. The researcher would answer the questions and explain the points that need clarity. Before employing the questionnaires, a pilot study was conducted with some in-service ELT teachers to figure out how well it works in the authentic context. The pilot study was conducted in the spring term of 2020-2021 academic year. While participants were answering the questionnaires, they were asked to read the instructions and items carefully and state the items that they do not comprehend. The instrument was revised based on the possible feedback from the pilot study. Following this, the main study was carried out in the fall term of 2021-2022 academic year.

As for the interviews, it was piloted with two in-service EFL teachers prior to having interviews, which would contribute to control the format and length of the questions and to reflect on the researcher's interviewing skills. This enabled the researcher to be aware of whether there are some unclear points or problems regarding the wording of the questions. This pilot study was also recorded with the aim of enhancing interviewing skills. After carrying out the pilot study, the interviews were conducted with 20 in-service EFL teachers through phone. To schedule the interviews, they were asked to pick suitable times for themselves. Accordingly, the interview schedule was organized as can be seen in Appendix D.

In the beginning, the interviewees were given information about the aim of the interview in relation to the aim of the study and ensured that they would be given pseudonyms in reporting the results of analysis. As the participants were guaranteed confidentiality, audio recording was utilized for the interviews. When the researcher posed the interview questions, the participants weren't interrupted so that they could give relevant anecdotes and provide further comments. Upon the completion of interviews that took about 20-25 minutes, all interviews were transcribed.

Prior to delving into the specific data analysis procedures for each data type, it is important to discuss matters pertaining to reliability and validity. This study incorporated

both quantitative and qualitative components and employed various approaches to assess research standards. The quantitative aspect involved the utilization of validity and reliability measures, following the guidelines proposed by Creswell (2012). The qualitative component, on the other hand, focused on evaluating credibility and trustworthiness, as recommended by Saldaña and Omasta (2018).

Questionnaire: Reliability and Validity

Reliability refers to the consistency and dependability of results obtained from a data collection tool (Creswell, 2012). It signifies that if the same data collection tool is administered multiple times and at different intervals, the results should closely resemble one another. Reliability is also associated with the consistency of responses given by individuals to related questions (Creswell, 2012, p.159).

Various types of reliability are defined in the literature, including test-retest reliability, alternate forms reliability, alternate forms and test-retest reliability, interrater reliability, and internal consistency (Creswell, 2012). Test-retest reliability involves administering the same data collection tool to participants at different times. However, in the present study, test-retest reliability was not utilized due to participant selection and sampling procedures. Alternate forms reliability requires the use of similar data collection tools with the same content, difficulty, or types. This type of reliability was not evaluated in this study, considering its aims and the specific data collection tools used. The third type combines the previous two types and involves administering the same type of test twice at separate times, which was also not applicable to this study for the reasons mentioned earlier.

Interrater reliability is relevant when the study involves observations of behavior, where different observers participate and compare their observation scores for reliability (Creswell, 2012, p.161). Intercoder reliability, a form of interrater reliability, refers to the degree of agreement between different coders on the same coding structure (Lavrakas, 2008). Since intercoder reliability pertains to qualitative data, further details will be discussed in the subsequent section.

Internal consistency was also a concern in this study, examining whether a participant's scores remained consistent across the items in the data collection tool. Consistency among scores indicates that the instrument's scores are reliable and accurate. To assess internal consistency, Cronbach's coefficient alpha was utilized (Creswell, 2012). It is argued that if the items are designed as continuous variables, such as in the case of a Likert scale (ranging from strongly agree to strongly disagree), Cronbach's alpha can be employed to test internal consistency. Additionally, Subedi (2016) suggested using Cronbach's alpha to test the reliability of Likert scales. Given that the questionnaire items in the present study are in Likert scale format, internal consistency evaluation procedures were followed.

To evaluate the reliability of the questionnaire, the items were grouped, and coefficient alpha scores were calculated for each scale. A score of .87 is considered satisfactory by Creswell (2012). A summary of the scores for each scale is provided in Table 8. The scale related to Knowledge of Testing and Assessment obtained the highest score (.86), while the lowest score was associated with Classroom-focused LTA (.80). Overall, the items within the scales can be considered reliable, and the obtained results can be deemed accurate.

Table 8

Questionnaire Parts and Reliability Scores

	Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
Classroom-focused LTA	.801	.801	36
Knowledge of Testing and Assessment	.855	.857	18
Purposes of Testing	.832	.832	14
Content and concepts of LTA	.844	.845	44
All parts	.874	.872	112

Another aspect to consider when creating and administering questionnaires is validity, which refers to the extent to which the available evidence supports the intended interpretation of test scores for the proposed purpose. In simpler terms, validity indicates the accuracy of the instrument used. Just like reliability, validity can be categorized into different types, such as content validity and criterion-related validity. Content validity assesses the extent to which a test measures the intended content area and requires both item validity and sampling validity. Item and sampling validity involve experts examining the items that make up the instrument. To ensure content validity, a pilot study was conducted with the target population, and interviews were conducted with both experts and participants to validate the items on the questionnaire.

In this particular research, a pilot study was carried out to enhance validity. As explained before, participants' responses were analyzed to identify any difficulties they encountered in understanding the items or completing the questionnaire. Additionally, interviews were conducted with the participants to confirm that the items effectively measured their intended constructs. Moreover, two university experts with PhD degrees in ELT were interviewed regarding the questionnaire. They confirmed that the items in the questionnaire were satisfactory and did not require any further corrections or editing.

Another type of validity is criterion-related validity, which assesses the correlation between scores on two different measures. It can be categorized into concurrent validity, where two separate tests are taken at the same time, and predictive validity, where the same test is taken multiple times. However, for this study, criterion-related validity was not considered since the data collection occurred only once.

Additionally, according to Creswell (2012), reliability and validity are closely connected. The source suggests that "these two terms sometimes overlap and at other times are mutually exclusive" (p.159). It further proposes that if the results obtained from a tool are reliable, it can be assumed that they are also valid. Keeping this in mind, the present

study takes both aspects into account and employs relevant measures to ensure both reliability and validity of the data collected.

Qualitative Data Collection Tool: Credibility and Trustworthiness

Lincoln and Guba (as cited in Saldaña & Omasta, 2018) argued that the standards used to evaluate quantitative data collection tools cannot be applied to qualitative research. Instead, they proposed the use of credibility and trustworthiness as alternative concepts to reliability and validity. Credibility refers to the plausibility of the researcher's work. To establish credibility, the researcher needs to provide a detailed description of the methodology, data collection tools, participants, and data analysis procedures. In this chapter, the research procedures are described in detail to ensure credibility. Trustworthiness, on the other hand, is about providing confidence in the research. It involves informing readers about the research procedures, such as the amount of qualitative data gathered, or the time spent in the field. In this study, trustworthiness is addressed by explaining the nature of the analysis and the work conducted within the scope of the study. The participants involved in interviews are described in detail while considering ethical concerns, the interview procedures are explained, and the amount of data gathered is stated. In essence, the research procedures are transparently presented to establish credibility and trustworthiness.

Furthermore, intercoder reliability is utilized to assess the level of agreement among different coders (Lavrakas, 2008). In this case, the coders analyze the same set of qualitative data, which consists of interviews, and then compare their codes to ensure consistency of interpretation (Saldaña & Omasta, 2018). Intercoder reliability is considered crucial for maintaining the quality of a research study (Johnson & Christensen, 2017). It is argued that intercoder reliability enhances objectivity and reduces biases that may arise from inconsistencies among coders (Johnson & Christensen, 2017).

To ensure intercoder reliability, two additional ELT graduates with PhD degrees were involved in analyzing the codes for the open-ended questions in the interview data.

The practice of involving multiple independent coders to ensure reliability is supported by existing literature (Barbour, 2001; Campbell, Quincy, Osserman & Pedersen, 2013) and the analyses show that intercoder reliability observed in this study can be deemed satisfactory.

Data Analysis

In this part, statistical procedures employed for the quantitative part and coding procedures utilized in the qualitative part of the current study are presented, respectively. For the data analysis of the quantitative data, descriptive and inferential statistics were employed. With the aim of determining the kind of inferential statistics test to be utilized, a test of normality was carried out through Kolmogorov-Smirnov and Shapiro-Wilk tests. The normality level of the questionnaire is shown in the table below.

Table 9

Normality Test of the Questionnaire

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	Df	Sig.	Statistic	df	Sig.
Questionnaire	.035	300	.200	.995	300	.447

a. Lilliefors Significance Correction

As presented in the table, both Kolmogorov-Smirnov ($p=.200 > .05$) and Shapiro-Wilk tests ($p=.447 > .05$) show that the data are normally distributed; therefore, parametric statistical tests could be utilized. As well as the data provided, normal Q-Q plot and histogram are shown in Figures 3 and 4.

Figure 3

Histogram for Normality Test of the Questionnaire

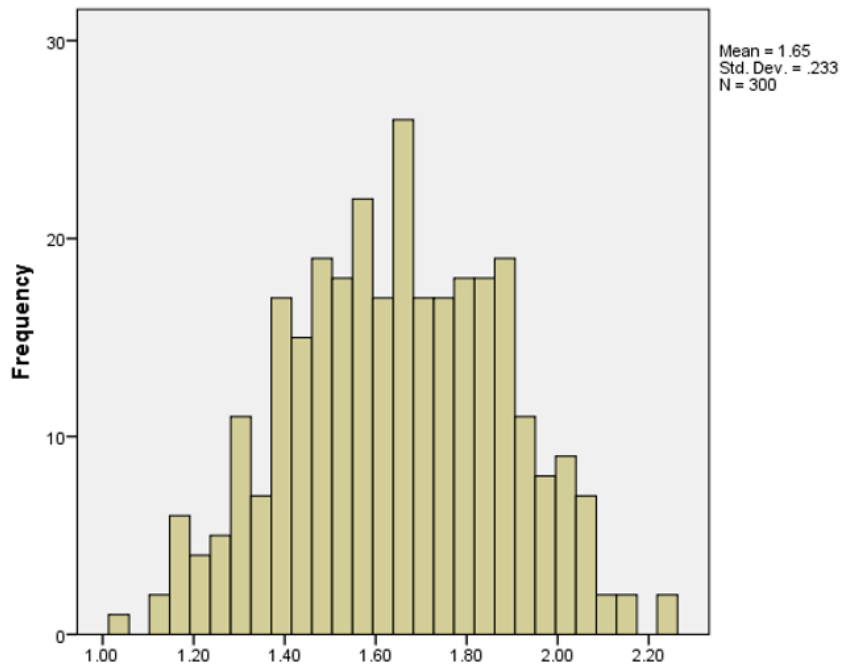
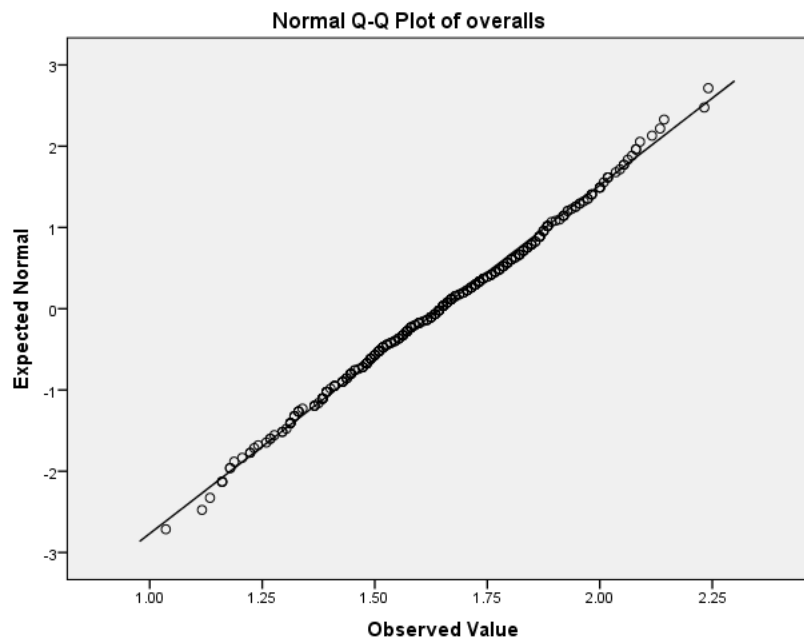


Figure 4

Normal Q-Q Plot for Normality Test of the Questionnaire



Figures 3 and 4 that present visual signs and support the data regarding normality reveal a normal distribution. After these, the quantitative data were analyzed.

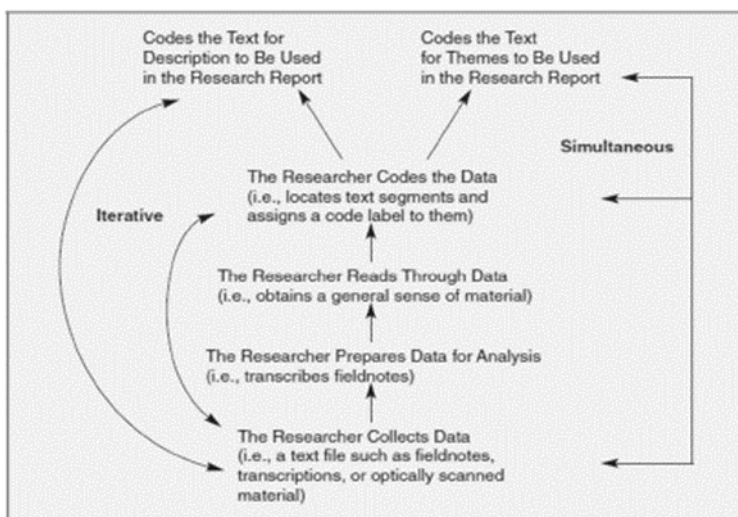
Regarding the first two research questions, they were analyzed by using descriptive statistics that include percentages, frequencies, mean values and standard deviation scores. In this regard, descriptive statistics would be applied to investigate the training that in-service language teachers receive and need by calculating mean values, frequencies, and percentages. With regard to the third research question, one-way between-groups analysis of variance (one-way ANOVA) was conducted to find out whether there are any statistically significant differences in training needs regarding LTA among in-service EFL teachers according to (a) the type of school/institution they teach at, (b) grade level they teach, (c) their level of experience, and (d) their education level. Moreover, for significant differences in findings, post hoc test (Tukey HSD) was employed in order to give more information about the variables which lead to important differences (Field, 2018).

As for the last research question, the analysis of the qualitative data was based on Grounded Theory (Strauss & Corbin, 1998) involving open and axial coding procedures. While open coding is the first step of analysis through which categories are specified, axial coding is related to conceptual ordering through which the categories and themes according to the prior coding process are identified. Following this, codes were classified to obtain wider categories and later themes (Cohen, Manion, & Morrison, 2018; Patton, 2002).

Based on the coding process, the data should be visited and revisited many times so that categories and themes can be reached for integrative grasp and analysis. Once the themes were identified, they were compared with the results of the questionnaire to establish connections between the qualitative and quantitative data, and to confirm the findings through triangulation. More specifically, this process to be followed is iterative and can be seen in Figure 5 below.

Figure 5

Qualitative Data Analysis Process (Creswell, 2012, p. 237)



Chapter 4

Findings and Discussion

RQ1: In-Service EFL Teachers' Received Training in Language Testing and Assessment

The first research question of the study was “How do in-service EFL teachers evaluate their undergraduate training in language testing and assessment? Do they perceive this training as adequate?”. This study aimed to explore how the in-service EFL teachers regard the training in LTA they underwent in their pre-service teacher education and whether this training is evaluated sufficient or not. There were 56 items in four sections which are Classroom-focused LTA, Knowledge of Testing and Assessment, Purposes of Testing, and Content and concepts of LTA. For this purpose, the respondents completed a four-point Likert scale (None; Little (1-2 days); Sufficient; Advanced). The findings of each item that reveal their training received in LTA domains are illustrated under the four categories.

Training Received in Classroom-focused LTA

This part consisted of 18 items overall. Table 10 below illustrates the findings related to the in-service EFL teachers' perceptions of the amount of training they received in classroom-focused LTA.

Table 10

Summary of the findings for training received in classroom-focused LTA

	N	Mean ^a	SD	None		Little (1-2 days)		Sufficient		Advanced	
				freq	%	freq	%	freq	%	freq	%
Preparing classroom tests	300	1.38	.835	54	18	93	31	138	46	15	5
Preparing diagnostic tests	300	1.26	.766	49	16.3	135	45	106	35.3	10	3.3
Preparing achievement tests	300	1.26	.804	51	17	138	46	94	31.3	17	5.7
Preparing proficiency tests	300	1.44	.763	30	10	128	42.7	122	40.7	20	6.7
Preparing placement tests	300	1.37	.825	46	15.3	119	39.7	114	38	21	7
Preparing progress tests	300	1.38	.811	41	13.7	126	42	111	37	22	7.3
Preparing language aptitude tests	300	1.51	.824	32	10.7	115	38.3	121	40.3	32	10.7
Using ready-made tests from textbook packages or from other sources	300	1.45	.814	35	11.7	123	41	115	38.3	27	9
Adapting ready-made tests for the needs of students	300	1.46	.802	29	9.7	134	44.7	107	35.7	30	10
Stages of language test construction	300	1.54	.819	26	8.7	123	41	114	38	37	12.3
Scoring	300	1.46	.815	33	11	125	41.7	113	37.7	29	9.7
Grading	300	1.44	.793	31	10.3	131	43.7	112	37.3	26	8.7
Giving feedback to students based on information from tests/assessment	300	1.30	.783	42	14	143	47.7	97	32.3	18	6
Interpreting test scores	300	1.20	.790	57	19	137	45.7	94	31.3	12	4
Using self/peer assessment	300	1.27	.848	57	19	126	42	96	32	21	7
Using informal, non-test type of assessment	300	1.37	.881	54	18	107	35.7	112	37.3	27	9
Using continuous type of assessment	300	1.43	.841	43	14.3	112	37.3	119	39.7	26	8.7
Using European Language Portfolio	300	1.56	.850	32	10.7	107	35.7	122	40.7	39	13

a.Means are based on a 4-point scale: 0, None; 1, Little (1-2 days); 2, Sufficient; 3, Advanced.

The overall mean score of training received in the area of Classroom-focused LTA is 1.39 (SD=.351), which reveals that in-service EFL teachers' training underwent in this area is little.

“Using European Language Portfolio” has the highest mean score (M=1.56, SD=.85) as the majority of them uttered that they received sufficient and little training (40.7%, 35.7%, respectively). “Stages of language test construction” has the second highest mean score (M=1.54, SD=.82). That is, nearly four in ten participants (41%) reported that the amount of training they had received was little. For the item “preparing language aptitude tests”, they reported sufficient level of training with the percentage of 40.3. Similarly, equally as many respondents reported that they perceived the training to be sufficient in “preparing proficiency tests” (40.7%). The number of the participants reporting no training received about “giving feedback to students based on information from tests/assessment” and “using continuous type of assessment” were also close to each other with the percentages of 14 and 14.3, respectively. However, “interpreting test scores” has the lowest mean score with 1.20 (SD=.79) since nearly one fifth of them (19%) reported no training in this constituent. “Preparing diagnostic” and “preparing achievement tests” followed the item the mean value of 1.26 (SD=.77 and .80, respectively) and almost 17% of them enunciated no training was offered to them. Also, “using self / peer assessment” had the the mean value of 1.27 (SD=.85) in that 19% of teachers stated no training taken. “Giving feedback to students based on information from tests/assessment” was the other aspect in which 14 % of them expressed lack of training.

Training Received in Knowledge of Testing and Assessment

There were 9 items in total. Table 11 below shows the findings regarding the in-service EFL teachers' perceptions of the amount of training taken in Knowledge of Testing and Assessment.

Table 11

Summary of the findings for training received in Knowledge of Testing and Assessment

		N	Mean ^a	SD	None		Little (1-2 days)		Sufficient		Advanced	
					freq	%	freq	%	freq	%	freq	%
Informal/ assessment	Formal	300	1.40	.877	53	17.7	101	33.7	120	40	26	8.7
Formative/ Summative assessment		300	1.59	.882	37	12.3	93	31	127	42.3	43	14.3
Norm /Criterion-referenced assessment		300	1.63	.910	40	13.3	79	26.3	132	44	49	16.3
Discrete point/Integrative testing		300	1.45	.943	56	18.7	94	31.3	110	36.7	40	13.3
Direct/Indirect testing		300	1.56	.885	36	12	104	34.7	116	38.7	44	14.7
Objective/Subjective testing		300	1.69	.878	32	10.7	81	27	136	45.3	51	17
Approaches to language testing		300	1.60	.874	35	11.7	94	31.3	128	42.7	43	14.3
Alternative assessment		300	1.55	.843	33	11	105	35	126	42	36	12
Computer-based testing		300	1.65	.810	28	9.3	84	28	152	50.7	36	12

a.Means are based on a 4-point scale: 0, None; 1, Little (1-2 days); 2, Sufficient; 3, Advanced.

The results of the training taken in this domain are illustrated with an overall mean score of 1.56 (SD=.536), indicating that the in-service language teachers reported the training they got in their undergraduate education to be nearly enough.

Of all the items related to Knowledge of Testing and Assessment, “objective / subjective testing” has the highest mean score (M=1.69, SD=.88) as 45.3% of them reported their training levels as enough. The second highest mean score was found for “computer-based testing” with the mean of 1.65 (SD=.81) and more than half of them (50.7%) expressed sufficient level of training in this item. The third highest mean score was identified for “norm / criterion-referenced assessment” (M=1.63, SD= .91) and the number

of participants stating adequate level of training was 44%. The number of respondents stating the training taken to be sufficient in “approaches to language testing” and “formative / summative assessment” were close to each other with 42.7% and 42.3%, respectively. Conversely, nearly one fifth of the respondents (17.7%) stated that they did not take any training in “informal / formal assessment”, which had the lowest mean score. The second lowest mean score was “discrete point / integrative testing” (M=1.45, SD= .94) as 18.7% of them reported no training.

Training Received in Purposes of Testing

This part consisted of 7 items overall. Table 12 below illustrates the findings related to the in-service EFL teachers' perceptions of the amount of training they received in the domain of Purposes of Testing.

Table 12

Summary of the findings for training received in Purposes of Testing

	N	Mean ^a	SD	None		Little (1-2 days)		Sufficient		Advanced	
				freq	%	freq	%	freq	%	freq	%
Giving grades	300	1.45	.937	56	18.7	91	30.3	115	38.3	38	12.7
Finding out what needs to be learned/taught	300	1.49	.973	62	20.7	71	23.7	125	41.7	42	14
Placing students onto programs, courses, etc.	300	1.57	.953	52	17.3	71	23.7	130	43.3	47	15.7
Testing competence in a language	300	1.52	.941	51	17	86	28.7	119	39.7	44	14.7
Identifying what has been learned	300	1.48	.969	56	18.7	90	30	107	35.7	47	15.7
Measuring general ability to learn a foreign language	300	1.56	.925	44	14.7	91	30.3	118	39.3	47	15.7
Awarding final certificates	300	1.55	.947	47	15.7	90	30	113	37.7	50	16.7

a.Means are based on a 4-point scale: 0, None; 1, Little (1-2 days); 2, Sufficient; 3, Advanced.

The results of the training taken in that domain are shown with an overall mean score of 1.51 (SD=.653), showing that the in-service language teachers reported the training they got in their undergraduate education to be slightly sufficient.

When the mean scores were compared to one another, it was found that items of this domain had closer mean values. That's, the range between the highest mean score and the lowest mean score were close to each other. "Placing students onto programs, courses, etc" has the highest mean score (M=1.57, SD=.95) as 43.3 % of them reported their training in this component was adequate for them. Similarly, equally as many respondents expressed that training in "measuring general ability to learn a foreign language" (39.3%) and "awarding final certificates" (37.7%) was adequate. As for the item of "giving grades" (M=1.45, SD=.94), 18.7% of them signified no training at all. Also, the same percentage of participants stated no training for "identifying what has been learned". (M=1.48, SD= .97).

Training Received in Content and Concepts of LTA

This domain involved 22 items. Table 13 below shows the findings of the in-service EFL teachers' perceptions of the amount of training taken in Content and concepts of LTA.

Table 13

Summary of the findings for training received in Content and Concepts of LTA

	N	Mean ^a	SD	None		Little (1-2 days)		Sufficient		Advanced	
				freq	%	freq	%	freq	%	freq	%
Testing reading in English	300	1.73	.965	32	10.7	94	31.3	96	32	78	26
Different test items/task types to test reading in English	300	1.85	.903	22	7.3	81	27	116	38.7	81	27
Testing listening in English	300	1.65	.947	37	12.3	94	31.3	106	35.3	63	21
Different test items/task types to test listening in English	300	1.68	.873	26	8.7	100	33.3	119	39.7	55	18.3

Testing speaking in English	300	1.46	.958	54	18	100	33.3	100	33.3	46	15.3
Different test items/task types to test speaking in English	300	1.69	.965	36	12	93	31	100	33.3	71	23.7
Testing writing in English	300	1.62	.915	33	11	104	34.7	106	35.3	57	19
Different test items/task types to test writing in English	300	1.63	.964	40	13.3	94	31.3	102	34	64	21.3
Testing Grammar in English	300	1.65	.937	32	10.7	106	35.3	97	32.3	65	21.7
Different test items/task types to test grammar in English	300	1.74	.987	32	10.7	98	32.7	85	28.3	85	28.3
Testing Vocabulary in English	300	1.64	.969	37	12.3	102	34	92	30.7	69	23
Different test items/task types to test vocabulary in English	300	1.67	.944	33	11	101	33.7	99	33	67	22.3
Testing integrated language skills	300	1.64	.859	24	8	111	37	113	37.7	52	17.3
Testing pronunciation in English	300	1.63	.940	34	11.3	106	35.3	97	32.3	63	21
Different test items/question types to test pronunciation in English	300	1.88	.886	15	5	92	30.7	106	35.3	87	29
Practicality	300	1.89	.911	20	6.7	83	27.7	108	36	89	29.7
Reliability	300	1.78	.922	29	9.7	81	27	118	39.3	72	24
Validity	300	1.84	.977	31	10.3	77	25.7	100	33.3	92	30.7
Authenticity	300	1.97	.943	24	8	66	22	106	35.3	104	34.7
Washback	300	1.86	.987	32	10.7	73	24.3	99	33	96	32
Using statistics to study the quality of tests / assessment	300	1.90	.898	18	6	83	27.7	110	36.7	89	29.7
Alternatives in assessment	300	1.73	.990	39	13	81	27	102	34	78	26

a. Means are based on a 4-point scale: 0, None; 1, Little (1-2 days); 2, Sufficient; 3, Advanced.

The overall mean score of training got in this field is 1.73 (SD=.42), which reveals that in-service EFL teachers' training taken in this area is nearly satisfactory.

“Authenticity” has the highest mean score, which is 1.97. (SD=.943) as 35.3% and 34.7% of in-service EFL teachers uttered that they received sufficient and advanced amounts of training, respectively. “Using statistics to study the quality of tests / assessment” follows with the mean value of 1.9 (SD=.90) since the majority specified either sufficient (36.7%) or advanced (29.7%) level of training during their undergraduate education. Similarly, (M=1.89, SD=.91) the majority of respondents expressed adequate and advanced levels of training for “practicality” with the percentages of 36 and 29.7, respectively.

Moreover, 35.3% and 29% of them stated that they received adequate and advanced level of training in “different test items/question types to test pronunciation in English”, respectively (M=1.88, SD=.89). For “washback” (M=1.86, SD=.99), 33% and 32% of them expressed sufficient and high level of training, respectively. In the same vein, “different test items/task types to test reading in English” has the mean value of 1.85 (SD=.90) as more than half of them (38.7% and 27%, respectively) revealed adequate and advanced level of training. “Regarding “validity” (M=1.84, SD=.98), 33.3% and 30.7% of in-service EFL teachers stated sufficient and advanced level of training.

As for the other items including testing writing in English (M=1.62, SD=.92), testing pronunciation in English (M=1.63, SD=.94), different test items/task types to test writing in English (M=1.63, SD=.96), testing integrated language skills (M=1.64, SD=.86), testing vocabulary in English (M= 1.64, SD=.97), testing grammar in English (M=1.65, SD=.94) and testing listening in English (M=1.65, SD=.95), their mean scores were so close to one another.

Conversely, nearly one fifth of the respondents (18%) stated that they did not take any training in “testing speaking in English”, which had the lowest mean score (M=1.46, SD=.958).

Training Received in Four Areas of LTA

In this part, there were 56 items and in-service EFL teachers responded to these items. Their answers were analyzed to answer the first research question. Table 14 below illustrates the findings of training taken in four areas of LTA.

Table 14

Summary of the findings for training received in four areas of LTA

Areas of LTA	N	Mean ^a	SD
Classroom-focused LTA	300	1.3931	.35138
Knowledge of Testing and Assessment	300	1.5678	.53674
Purposes of Testing	300	1.5186	.65355
Content and Concepts of LTA	300	1.7338	.42044
Total	300	1.5707	.26480

a.Means are based on a 4-point scale: 0, None; 1, Little (1-2 days); 2, Sufficient; 3, Advanced.

The overall mean score of these four domains was 1.57, revealing that in-service EFL teachers expressed that their training in these domains during their university education was almost inadequate, though not totally. Among all the domains of LTA, “content and concepts of LTA” has the highest mean score (M=1.73, SD=.42). The second highest mean score was the domain of “knowledge of testing and assessment” (M=1.56, SD=.53) and it was followed by the domain of “purposes of testing” (M=1.51, SD=.65). “Classroom-focused LTA” has the lowest mean value with 1.39 (SD=.35).

Overall, they perceived the training they took in the domain of “content and concepts of LTA” was more than other three domains and the domain of “classroom-focused LTA” was found to be less than the others.

RQ2: EFL Teachers' In-service Training Needs in Language Testing and Assessment

The second research question of the study was “To what extent do in-service EFL teachers perceive a need for in-service training in language testing and assessment?”. With this question, the study tried to identify the in-service EFL teachers' future training needs with regard to LTA field. To this end, four sections were identified which are Classroom-focused LTA, Knowledge of Testing and Assessment, Purposes of Testing, and Content and concepts of LTA. In total, there were 56 items, and the participants completed a four-point Likert scale (None; Basic training; Intermediate training; Advanced training). The findings of each item that show participants' perceived LAL in-service training needs are shown under these four categories.

Training Needs in Classroom-focused LTA

In this part, there were 18 items in total. Table 15 below illustrates the findings related to the in-service EFL teachers' perceptions of the amount of training needed in classroom-focused LTA.

Table 15

Summary of the findings for in-service training needs in classroom-focused LTA

	N	Mean ^a	SD	None		Basic training		Intermediate training		Advanced training	
				freq	%	freq	%	freq	%	freq	%
Preparing classroom tests	300	1.43	.895	49	16.3	108	36	109	36.3	34	11.3
Preparing diagnostic tests	300	1.51	.848	34	11.3	116	38.7	114	38	36	12
Preparing achievement tests	300	1.56	.850	32	10.7	108	36	121	40.3	39	13
Preparing proficiency tests	300	1.52	.795	28	9.3	118	39.3	125	41.7	29	9.7
Preparing placement tests	300	1.54	.843	31	10.3	113	37.7	118	39.3	38	12.7
Preparing progress tests	300	1.51	.844	34	11.3	114	38	117	39	35	11.7

Preparing language aptitude tests	300	1.44	.810	37	12.3	117	39	122	40.7	24	8
Using ready-made tests from textbook packages or from other sources	300	1.35	.831	49	16.3	117	39	114	38	20	6.7
Adapting ready-made tests for the needs of students	300	1.30	.824	53	17.7	121	40.3	109	36.3	17	5.7
Stages of language test construction	300	1.35	.763	41	13.7	126	42	121	40.3	12	4
Scoring	300	1.31	.788	50	16.7	119	39.7	120	40	11	3.7
Grading	300	1.26	.796	53	17.7	130	43.3	104	34.7	13	4.3
Giving feedback to students based on information from tests/assessment	300	1.33	.802	48	16	120	40	117	39	15	5
Interpreting test scores	300	1.46	.786	33	11	117	39	128	42.7	22	7.3
Using self/peer assessment	300	1.51	.791	31	10.3	110	36.7	134	44.7	25	8.3
Using informal, non-test type of assessment	300	1.45	.858	46	15.3	101	33.7	126	42	27	9
Using continuous type of assessment	300	1.45	.900	48	16	104	34.7	112	37.3	36	12
Using European Language Portfolio	300	1.50	.867	40	13.3	105	35	120	40	35	11.7

a.Means are based on a 4-point scale: 0, None; 1, Basic training; 2, Intermediate training; 3, Advanced training.

The overall mean score of training needed in the area of Classroom-focused LTA is 1.43 (SD=.386), which reveals that in-service EFL teachers' further training need in this area is basic.

Of all the items related to classroom-focused LTA, "preparing achievement tests" has the highest mean score (M=1.56, SD=.85). That's, 40.3 and 36% of in-service EFL teachers reported that they needed intermediate and basic training, respectively. The second highest mean score was identified for "preparing placement tests" (M=1.54, SD=.84) since 39.3 and 37.7% of the participants revealed intermediate and basic training,

respectively. This was followed by “preparing proficiency tests” (M=1.52, SD=.79), “preparing progress tests” (M=1.51, SD=.84), “preparing diagnostic tests” (M=1.51, SD=.85), “using self/peer assessment” (M=1.51, SD=.79). For the item “using European Language Portfolio”, they reported intermediate and advanced level of training with the percentages of 40 and 11.7, respectively. On the other hand, “grading” has the lowest mean score with 1.26 (SD=.79) and “adapting ready-made tests for the needs of students” has the second lowest mean score with 1.30 (SD=.82) since 17.7% of the participants stated no further training need in these two items. It was followed by “scoring” (M=1.31, SD=.79) as 16.7% of the in-service EFL teachers did not need any more training in this item.

Training Needs in Knowledge of Testing and Assessment

There were 9 items in total in this part. Table 16 below illustrates the findings related to the in-service EFL teachers' perceptions of the amount of training needed in Knowledge of Testing and Assessment.

Table 16

Summary of the findings for in-service training needs in Knowledge of Testing and Assessment

		N	Mean ^a	SD	None		Basic training		Intermediate training		Advanced training	
					freq	%	freq	%	freq	%	freq	%
Informal/assessment	Formal	300	1.71	.788	21	7	86	28.7	152	50.7	41	13.7
Formative/Summative assessment		300	1.72	.814	21	7	89	29.7	142	47.3	48	16
Norm-referenced assessment	/Criterion-referenced	300	1.64	.883	34	11.3	88	29.3	130	43.3	48	16
Discrete point/testing	Integrative testing	300	1.64	.872	30	10	98	32.7	123	41	49	16.3
Direct/testing	Indirect testing	300	1.62	.898	35	11.7	95	31.7	120	40	50	16.7
Objective/testing	Subjective testing	300	1.66	.935	36	12	92	30.7	111	37	61	20.3

Approaches to language testing	300	1.56	.907	37	12.3	107	35.7	107	35.7	49	16.3
Alternative assessment	300	1.73	.903	25	8.3	99	33	109	36.3	67	22.3
Computer-based testing	300	1.62	.972	44	14.7	87	29	107	35.7	62	20.7

a.Means are based on a 4-point scale: 0, None; 1, Basic training; 2, Intermediate training; 3, Advanced training.

Descriptive statistics shows the results of the further training stated in the area of Knowledge of Testing and Assessment with an overall mean score of 1.65 (SD=.56), indicating that the in-service EFL teachers perceive a need for further intermediate training in this field of LTA.

“Alternative assessment” has the highest mean score (M=1.73, SD=.903). This reveals that more than one third (36.3%) of in-service EFL teachers uttered intermediate training need in this item. “Formative / summative assessment” follows it with 1.72 mean score (SD=.81) as nearly half of them (47.3%) stated training need at intermediate level. It was followed by “informal / formal assessment” (M=1.71, SD=.79) as 50.7% of them stated training need that is intermediate level. For the item “objective / subjective testing”, they reported intermediate and advanced level of training with the percentages of 37 and 20.3, respectively. As for “norm / criterion-referenced assessment”, they stated reported intermediate and advanced level of training with the percentages of 43.3 and 16, respectively. With regard to the lowest mean score, it was found for “approaches to language testing” (M=1.56, SD=.90) since nearly one in ten (12.3%) of them did not state any further training need. It was followed by two components which are “direct / indirect testing” (M=1.62, SD= .90) and “computer-based testing” (M=1.62, SD=.97). In-service EFL teachers reported no need for training on “direct / indirect testing” (11.7%) and for “computer-based testing” (14.7%).

Training Needs in Purposes of Testing

In this part, there were 7 items in total. Table 17 below illustrates the findings related to the in-service EFL teachers' perceptions of the amount of training needed in Purposes of Testing.

Table 17

Summary of the findings for in-service training needs in Purposes of Testing

	N	Mean ^a	SD	None		Basic training		Intermediate training		Advanced training	
				freq	%	freq	%	freq	%	freq	%
Giving grades	300	1.44	.888	51	17	96	32	123	41	30	10
Finding out what needs to be learned/taught	300	1.71	.936	35	11.7	83	27.7	117	39	65	21.7
Placing students onto programs, courses, etc.	300	1.72	.982	39	13	81	27	105	35	75	25
Testing competence in a language	300	1.88	.974	30	10	72	24	102	34	96	32
Identifying what has been learned	300	2.16	.906	18	6	49	16.3	101	33.7	132	44
Measuring general ability to learn a foreign language	300	2.10	.914	20	6.7	51	17	107	35.7	122	40.7
Awarding final certificates	300	2.03	.993	28	9.3	59	19.7	89	29.7	124	41.3

a.Means are based on a 4-point scale: 0, None; 1, Basic training; 2, Intermediate training; 3, Advanced training.

The overall mean score of training needed in the field of Purposes of Testing is 1.86 (SD=.57), which reveals that in-service EFL teachers' further training need in this domain is intermediate level.

Of all the items related to this domain, "identifying what has been learned" has the highest mean score (M=2.16, SD=.906). That's, 44 and 33.7% of in-service EFL teachers reported that they needed advanced and intermediate training, respectively. The second highest mean score was identified for "measuring general ability to learn a foreign language" (M=2.1, SD=.91) since in-service EFL teachers reported advanced and intermediate level of further training with the percentages of 40.7 and 35.7, respectively. For "awarding final

certificates”, 41.3% of them uttered advanced training need and 29.7% of them stated intermediate level training. As for “testing competence in a language”, the majority of them stated training need at different levels and only 10% of them stated they did not need any training. Regarding “placing students onto programs, courses, etc.”, they reported training needs as intermediate and advanced with the percentages of 35 and 25, respectively. Similarly, equally as many respondents reported that they perceived either intermediate (39%) or advanced training need (21.7%) related to “finding out what needs to be learned/taught”. On the other hand, “giving grades” has the lowest mean score with 1.44 (SD=.89) since 17% of the participants stated no further training need in this item.

Training Needs in Content and concepts of LTA

This part included 22 items. Table 18 below shows the findings pertaining to the in-service EFL teachers' perceptions of the amount of training needed in Content and concepts of LTA.

Table 18

Summary of the findings for in-service training needs in Content and concepts of LTA

	N	Mean ^a	SD	None		Basic training		Intermediate training		Advanced training	
				freq	%	freq	%	freq	%	freq	%
Testing reading in English	300	1.87	.962	29	9.7	74	24.7	105	35	92	30.7
Different test items/task types to test reading in English	300	1.82	1.008	32	10.7	87	29	83	27.7	98	32.7
Testing listening in English	300	1.85	1.047	40	13.3	69	23	86	28.7	105	35
Different test items/task types to test listening in English	300	1.81	.929	28	9.3	79	26.3	115	38.3	78	26
Testing speaking in English	300	1.81	.968	32	10.7	78	26	105	35	85	28.3
Different test items/task types to test speaking in English	300	1.66	.977	38	12.7	98	32.7	93	31	71	23.7
Testing writing in English	300	1.69	.995	35	11.7	104	34.7	80	26.7	81	27
Different test items/task types to test writing in English	300	1.66	.963	37	12.3	95	31.7	100	33.3	68	22.7
Testing Grammar in English	300	1.86	.927	21	7	90	30	100	33.3	89	29.7
Different test items/task types to test grammar in English	300	1.94	.934	25	8.3	65	21.7	113	37.7	97	32.3
Testing Vocabulary in English	300	1.95	.956	28	9.3	61	20.3	110	36.7	101	33.7
Different test items/task types to test vocabulary in English	300	2.05	.913	23	7.7	49	16.3	119	39.7	109	36.3
Testing integrated language skills	300	2.14	.970	30	10	33	11	103	34.3	134	44.7
Testing pronunciation in English	300	2.12	.957	24	8	49	16.3	94	31.3	133	44.3
Different test items/question types to test pronunciation in English	300	2.08	1.002	30	10	49	16.3	88	29.3	133	44.3
Practicality	300	2.14	.924	20	6.7	50	16.7	99	33	131	43.7

Reliability	300	2.14	.986	27	9	46	15.3	84	28	143	47.7
Validity	300	2.14	.947	22	7.3	50	16.7	92	30.7	136	45.3
Authenticity	300	2.04	.958	24	8	59	19.7	97	32.3	120	40
Washback	300	1.98	1.080	47	15.7	37	12.3	92	30.7	124	41.3
Using statistics to study the quality of tests / assessment Alternatives in assessment	300	1.95	1.000	32	10.7	62	20.7	96	32	110	36.7
	300	1.95	.972	30	10	59	19.7	107	35.7	104	34.7

a.Means are based on a 4-point scale: 0, None; 1, Basic training; 2, Intermediate training; 3, Advanced training.

The overall mean score of training needed in the field of Content and Concepts of LTA is 1.94 (SD=.384), which reveals that in-service EFL teachers' further training need in this domain is intermediate level.

The items of "testing integrated language skills", "practicality", "reliability", "validity" have the highest same mean score, which is 2.14. For "testing integrated language skills" (SD=.97), the majority of respondents stated advanced or intermediate levels of training (44.7% and 34.3%, respectively). Regarding "reliability", "validity", "practicality", equally as many reported that they perceived either advanced (47.7%, 45.3%, 43.7%, respectively) or intermediate training need (28%, 30.7%, 33%, respectively). It was followed by "testing pronunciation in English" (M=2.12, SD=.96) with a close mean value as they uttered the need for advanced and intermediate training with the percentages of 44.3 and 31.3, respectively. For "different test items/task types to test vocabulary in English" and "different test items/question types to test pronunciation in English" components, most of the participants pronounced the need for extra training and only 7.7% and 10% of them, stated no further training need respectively. The number of respondents stating advanced training need in "washback" and "authenticity" were also high and very close to each other with 41.3% and 40%, respectively. Also, the number of the participants reporting advanced training needed in "testing listening in English", "testing vocabulary in English" and "different

test items/task types to test grammar in English” were high and very close to one another with the percentages of 35, 33.7 and 32.3, respectively.

As to the lowest mean score, “different test items/task types to test speaking in English” and “different test items/task types to test writing in English” have the lowest same mean score, which is 1.66. 12.7% and 12.3% of the respondents did not express any training need for “different test items/task types to test speaking in English” and “different test items/task types to test writing in English”, respectively.

Training Needs in Four Areas of LTA

This part included 56 items and in-service EFL teachers’ responses to these items were analyzed to answer the second research question. Table 19 below illustrates the findings pertaining to the in-service EFL teachers’ perceptions of the amount of training needed in four domains of LTA.

Table 19

Summary of the findings for in-service training needs in four areas of LTA

Areas of LTA	N	Mean ^a	SD
Classroom-focused LTA	300	1.4315	.38639
Knowledge of Testing and Assessment	300	1.6548	.56057
Purposes of Testing	300	1.8624	.57087
Content and Concepts of LTA	300	1.9380	.38429
Total	300	1.7202	.26392

a.Means are based on a 4-point scale: 0, None; 1, Basic training; 2, Intermediate training; 3, Advanced training.

The overall mean score of these four domains was 1.72, revealing that in-service EFL teachers’ further training need in these domains is intermediate level. Of all the

domains of LTA, “content and concepts of LTA” has the highest mean score ($M=1.94$, $SD=.384$). The second highest mean score was the domain of “purposes of testing” ($M=1.86$, $SD=.57$) and it was followed by the domain of “knowledge of testing and assessment” ($M=1.65$, $SD=.56$). “Classroom-focused LTA” has the lowest mean value with 1.43 ($SD=.386$).

Overall, the need for further training perceived and uttered by in-service EFL teachers in the domain of “content and concepts of LTA” was more than other tree domains and need for further training in “Classroom-focused LTA” was found to be less than others.

RQ3: Differences between EFL Teachers’ Training Needs

The third research question of the study was the following: “Do in-service EFL teachers’ training needs in language testing and assessment vary according to different educational settings, the grade level they are currently teaching, years of experience and their education level?”. To answer this question, the participants’ educational settings they are working, grade level they are teaching, years of professional experience, and education level were compared by using one-way ANOVA.

Educational Setting

As mentioned earlier in the Methodology chapter, the participants were 300 in-service EFL teachers working in different formal and non-formal education settings in Turkey. One of the sub-aims was to determine if there were any significant differences between the participants’ training needs and educational settings. The results of one-way ANOVA test showed that there were no significant differences between participants’ further training needs in LTA with regards to education settings ($F(2,297) = .708$, $p = .493$).

Table 20

Summary of the ANOVA results for educational settings

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.099	2	.049	.708	.493
Within Groups	20.728	297	.070		
Total	20.827	299			

Grade Level of Teaching

In-service EFL teachers working at primary, secondary and high schools as English teachers attended the current study and so as to determine if their further training needs in LTA altered significantly based on the grade levels of teaching, one-way ANOVA test was utilized. The results showed that there were no significant differences between teachers working for different grade levels for further LTA training needs ($F(2,297) = .442, p = .643$).

Table 21

Summary of the ANOVA results for grade level of teaching

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.062	2	.031	.442	.643
Within Groups	20.765	297	.070		
Total	20.827	299			

Years of Experience

To determine if there were any significant differences between EFL teachers with different years of experience in their profession with regards to their further training needs, the same statistical procedures were followed. There were 95 EFL teachers with 1-5 years of experience, 112 teachers with 6-10 years of experience and 93 teachers with 11 or more years of experience in this study. The results of one-way ANOVA test presented that between years of teaching experience and training needs in LTA ($F(2,297) = 1.848, p = .159$), there were no significant differences ($p > .05$). To put it differently, again, teachers' years of experience was not an important factor that determines the future training needs of EFL teachers.

Table 22

Summary of the ANOVA results for years of experience

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.256	2	.128	1.848	.159
Within Groups	20.571	297	.069		
Total	20.827	299			

Education Level

Another condition for identifying further needs differences between EFL teachers was their education level. There were 27 participants with MA and 12 with PhD degrees while 261 of them held BA degree in the study. Once again, one-way ANOVA was administered to explore the differences between groups for their further LTA training needs. As the results showed, there were significant differences among teachers with different

education levels in terms of their further training needs in LTA ($F(2,297) = 3.542, p = .030$). In other words, the EFL teachers' training needs differed based on their level of education.

Table 23

Summary of the ANOVA results for level of education

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.485	2	.243	3.542	.030
Within Groups	20.342	297	.068		
Total	20.827	299			

Post hoc comparisons using the Tukey HSD test which was employed to identify the significant variables indicated that the mean score for the EFL teachers with MA degree ($M=1.85, SD=.28$) was significantly different from teachers with BA degree ($M=1.70, SD=.26$) at $p=.025$ level. However, there was no significance between other groups of teachers with regards to their training needs and education levels ($p > .05$).

Table 24

Summary of the Tukey HSD results for level of education

(A) Education Level	(B) Education Level	Mean Difference (A-B)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
BA	MA	-.13848*	.05291	.025	-.2631	-.0139
	PhD	.02422	.07727	.947	-.1578	.2062
MA	BA	.13848*	.05291	.025	.0139	.2631
	PhD	.16270	.09080	.174	-.0512	.3766
PhD	BA	-.02422	.07727	.947	-.2062	.1578
	MA	-.16270	.09080	.174	-.3766	.0512

*. The mean difference is significant at the 0.05 level.

To put it differently, the further training needs of EFL in-service teachers showed significant changes between holding MA and BA degrees. A summary of the results of Tukey test can be viewed in Table 24 above.

Results for RQ 4

As mentioned earlier, upon analyzing the data gathered through questionnaire, 20 voluntary participants were invited to take part in the qualitative part of the study. Out of 20 participants, 14 of them hold a BA degree (10 and 4, respectively) in ELT and ELIT, and 6 of them have an MA degree (3 and 3, respectively) in ELT and Curriculum and Instruction. There is no teacher with a PhD degree. Their teaching experience varies from 1-5 years to 11 or more years in that while 6 of them 1-5 years of experience and 5 of them have 6-10 years of teaching experience, 9 of them have 11 or more years in their profession. They

work at different schools that are public, private and private language schools in different teaching levels which are primary, secondary and high school.

As to the number of LTA and statistics courses taken in undergraduate education, all participants stated that they have taken 1 LTA course and only 3 of them expressed 1 statistics course in addition to LTA course.

Do you think that Testing and Evaluation Course that you took in undergraduate training covers necessary components of language testing and assessment?

When participants were addressed the question of adequateness of constituents involved in Testing and Evaluation Course, 5 of them specified that it was enough whereas 15 of them noted that it was insufficient for them.

“I think the training in LTA was enough for us as EFL teachers. As far as I remember we learned how to prepare and score tests as well as giving feedback. I think those are necessary and we could learn the other details in the profession.” (Selin, Private Primary School)

“For me, we learned some important topics in our university education. We took many methodology and literature courses which are so important for us. Also, we took 1 testing course I guess that is enough for me because I can use tests easily. I do not need specific course for this.” (Mustafa, Private Primary School)

These two teachers defended the idea that the course offered in their BA programs was adequate for including essential topics of LTA and they are not in need of learning more about this. Especially, Selin teacher indicated that teachers could discover something about testing as they work. Similarly, Mustafa teacher underlined the idea that he could use tests without much effort and expertise.

Moreover, one of the participants said that the training was enough to include essentials of testing and viewed their professors to be competent at this field by covering theoretical and practical aspects.

“I guess I know all the important concepts and terms and strategies in assessment thanks to my training in university. Our professors were expert in this area, and they taught us theory and practice of testing in a detailed way. For instance, I learned how to use statistics for testing purpose, which I find very useful.” (Ada, Public High School)

On the other hand, other in-service EFL teachers expressed that Testing and Evaluation Course offered in their BA's did not cover the required items. Most of them stated they realized this necessity in the profession.

“In the beginning, when I graduated, I did not think about this. But, when years passed, I noticed I missed something: assessment. I was not expert in this field, of course. But, I have to be better at least, I thought.” (Okan, Private Primary School)

“As a teacher, we had to receive more training because the course I took was not sufficient for me. I feel I know limited knowledge. Sometimes, when I talk to my other colleagues and I conduct testing, I realize the need for more information in testing to be creative and more motivating.” (Merve, Private Language School, Primary school)

“When I become a teacher, I think I am provided with all knowledge and practice. But now, I see that I need more in every field. One of them is testing, so I can easily say that the training is not enough. I guess almost all teachers experience the same thing.” (Eda, Public Secondary School)

“Actually, no it did not cover much. Testing is a huge area and includes many things. It is not only about testing students but also giving feedback and contributing to them through our observations and notes. I wish we had taken more courses to cover all of these.” (Dođa, Public Secondary School)

Of the topics covered in this course, which topic/s do you think is/are the MOST helpful to you as a classroom teacher?

In-service ELT teachers uttered that certain topics are the most fruitful to them as language teachers. When the topics itemized by the participants were analyzed, it was found that 8 components were listed. Alternative assessment, validity, how to adapt prepared tests, practicality and informal/formal assessment are stated. As well as these topics, objective testing, formative/summative assessment and testing reading are pronounced by them. They stated that these topics were so beneficial for them in their actual in-class experience.

“Keeping up with new generation of students requires different teaching and assessment knowledge and strategies. Acquainting myself with alternative assessment is necessary, I believe. Rather than learning to prepare classical ways of testing, I feel lucky that I know other ways like portfolios, presentations and self-assessment because they are more student- and process- focused, so as a teacher, alternative assessment is the most useful one for me to make them more concentrated.” (Seda, Private Secondary School)

“According to me, knowing informal and formal assessment is the most useful thing to be honest because we teach to Gen Z. If I had not learned these two, I would have probably used only formal assessment like quizzes and exams. But now, I use

observations and surveys, too. Fortunately, I learned these.” (Dilan, Private Primary School)

As can be understood from Seda and Dilan teacher’s statements, they think receiving training in certain aspects of assessment like alternative assessment and informal/formal assessment in pre-service teacher education are beneficial for them. They highlight the changing necessities of teaching to and assessing new generations of students. Particularly, Seda teacher emphasized the importance of capturing Gen Z students’ interests and motivating them through alternative assessment.

“Validity is important in testing and the first thing that comes into my mind is validity when I am asked the most useful topic because it is about whether the test measures what it aims to measure. Without being aware of this important concept, we cannot prepare or use any test, I strongly believe.” (Alican, Private Language School, Secondary School)

“As a teacher, practicality is crucial and the most helpful for me among the topics that I learned in my undergraduate education because I have a heavy teaching load and while I am preparing and scoring a test, I want it to be easy in all means. If I had to spend much time in preparing or scoring it, I could not focus on other tasks.” (Melis, Private High School)

The teachers stated that principles of assessment that include validity and practicality are fruitful for them on the ground that they have to be good at measuring what they seek to measure and preparing and using tests that are economical in time and effort.

“For me, I believe that objective testing is the most useful topic that I learned in my BA. Knowing the difference between objective and subjective testing is important

and I know which test types can be measured objectively like true / false questions and multiple-choice tests. Rather than evaluating my students according to my criteria, I prefer these types of tests and feel more secure.” (Berkin, Public Primary School)

“I can carry out formative and summative assessment thanks to my training in these. Without them, I may not know the purpose of these assessments like monitoring students during the process and evaluating them at the end of the course. I am lucky that I know these terms.” (Okan, Private Primary School)

Types of testing and functions of assessment are stated to be beneficial for teachers since they could use this understanding in their real classroom practice. Also, the topic Berkin teacher stated made him feel safer while the topic Okan teacher indicated made him feel happy.

“Testing reading is the most beneficial topic that we learn in undergraduate training, I believe. Based on the training I received, it is really important for learners to read something in English to develop their language overall and to assess them efficiently using the appropriate items. We learned different test items to measure reading skill like interactive and extensive reading. Those are so essential in a language class. I suggest that this subject continued to be given in teacher trainings. ” (Fatma, Public High School)

Fatma teacher emphasized the substantiality of utilizing different task types in reading rather than using standard tasks. She believed that training in this area should not be abandoned.

“Well, for me learning the ways to adapt prepared test is the most useful one because we are not experts in testing, and it is really difficult to prepare a reliable test. So, I do not trust in myself in preparing a test from scratch and I have to know the rules of adapting ready-made materials. For instance, simplifying the task or adjusting the length of texts.” (Cenk, Private Language School, Primary School)

Moreover, one of the participants stated that he is aware of the difficulties of preparing a qualified test and without having expertise in testing, he does not want to prepare it. Instead of this, thanks to having knowledge in adapting pre-existing materials of testing, he could adapt and employ the tests in his classrooms.

If you were to take this course again, or take an advanced course in language testing and assessment, what topics would you like to learn about?

During the interviews, when asked about their beliefs on the topics to learn, they stated various subjects that consist of statistics, alternative assessment, computer-based testing, reliability, approaches to language testing, giving feedback, testing integrated language skills and validity.

“When I think topics to learn more, two things come into my mind: computer-based testing and approaches to language testing. Currently, we need fast and easy systems to assess students and computer-based testing provides this, so I prefer learning more about it. As we all know, this generation requires different ways to teach and learn. Also, I heard from my colleagues that there are different approaches to testing, but I am not well-informed about them.” (Seher, Private Primary School)

“Let me think for a second... Well, it would be how to test different language skills at the same time. I could not remember the term of it now, but like combining language

skills like listening and writing. Especially, today's world requires many skills simultaneously and I have to prepare students for the real world. For example, they have to listen to something, take notes and speak at the same time." (Ahmet, Public Primary School)

"It is a difficult question to choose the topic but maybe rather than classical ways of tests and quizzes, we could learn other modern ways like observing and taking notes, keeping journals online etc... We need them more when we compare it to the past. Those would absolutely capture both us and students more." (Melis, Private school, High School)

These teachers asked for more training in the given areas above and they point to the need for these subjects when the needs of students and contemporary world are taken into account. Particularly, addressing to skills and abilities that are important and common in the world nowadays is mentioned.

"Recently, I hear about adaptive testing. I just know the major things about it like test items are listed based on students' level and correct answers. I want to discover how to use and design it to adapt to our real classrooms. Of course, we have to be good at technology, I mean we have to be able to use technology effectively." (Gizay, Private language school, High School)

Gizay teacher emphasized the need for using adaptive testing which is another form of computer-based testing. She states that she wants to employ this in her teaching context in that the system shows easier or more difficult questions depending on the correctness of examinee's responses.

"Definitely, it would be giving feedback. I am interested in different sorts of feedback. It is incredibly important for self-learning and self-awareness. This awareness can improve us. What I mean is ... I mean as a teacher we will also develop ourselves

through feedback and students will, too. As a language teacher, we have to learn different types.” (Seda, Private Secondary School)

“It is a thought-provoking question, I think... Well, probably if I would take a course in testing, I would like to learn validity and reliability. Maybe, one more thing: statistics. There are many types of tests both online and printed. But, I want to learn more about validity, for instance. I do not remember much, but there are few types. Also, I want to use more statistical ways to assess the students. It is better to learn them in detail both for our development and learning and students’.” (Alper, Private Language School, Secondary School)

As can be seen in Seda and Alper teachers’ responses, taking a course again on these subjects that they listed would serve as a development activity for their own learning and for students’. Seda teacher also underlined another term called “self-awareness” that is another contributing factor for development.

How do you decide which assessment method to use in your classrooms?

When in-service EFL teachers were addressed the question of the way they determine assessment methods to employ, they pointed out certain ways.

“I consult to my colleagues if I try to choose a method to assess students. First, I give information about my students and their strong as well as weak sides. Then, I reveal my aim and we try to find the best way to use in the classroom. Such collaboration is my first preference, and it is easier and quick.” (Gizay, Private language School, High School)

“Let me think... I guess I will ask other English teachers in the meetings or at other times. I decide in this way because it is practical and easy for me. Also, they know the school and classroom context that is an advantage for me.” (Seher, Private Primary School)

“Absolutely I talk to my colleagues because I mean we can learn from each other. Also, I am more comfortable with them when I need to decide on something including testing. I believe that cooperation among teachers is very essential. I mean... Without supporting each other, we cannot function truly.” (Okan, Private Primary School)

These teachers articulated that they decide on methods to utilize through asking their workmates. They viewed this way as practical and simpler. In this way, they emphasize the notion of cooperation. Furthermore, one of them stated that he feels more relaxed if he talks to one of his colleagues.

“For me, it is easy to use the internet for deciding. It has so many options and I can choose one of them according to my class. I do not put much effort thanks to the internet. Also, I see some creative ways to assess students, especially my students are younger, and I want to find different new methods so that they can be motivated even in the exam.” (Sezen, Public Primary School)

“I like searching about new ways and strategies to test. Otherwise, it would be the same and monotonous for me and students, too. I just entered the key words and lots of choices appear on the screen.” (Alper, Private Language School, Secondary school)

“I work with teenagers, and I like using different methods with them. There are many engaging methods on the net. I just spend 5 minutes and find something to use. That’s all.” (Doğa, Public Secondary School)

These teachers found using the internet to make their decisions on assessment methods undemanding and functional. They all stated that they could find diverse activities and assessment ideas to apply.

“Personally, I check theoretical books of testing just as I do for methodology purposes. I just open the related part and read it quickly, so I feel more secure and confident. I believe that is the best way.” (Fatma, Public High School)

“Let me think... Most of the time, I look at books on the issue. I keep my old books in different areas like teaching to young learners, approaches and methods, testing and literature. I take advantage of them. Sometimes, it sounds a bit difficult to find the books and scan them, but it is effective, I believe.” (Eda, Public Secondary school)

“I use my books to check anything about my profession and for testing, too. For instance, last week, I wanted to check something about testing. I guess it was summative assessment. I forgot about the definition. I wanted to see whether my goal and test type are compatible. So, I realized I was using it for formative purposes. Like this, I used this way for 7 years and it made me more confident” (Ayla, Public Secondary School)

As stated by some of the EFL teachers, they prefer checking books in testing and assessment to decide their methods or anything that they wondered. By doing so, they feel more confident.

“I try to remember what I learned in my testing courses. Sometimes, I have difficulty in remembering because, of course many years passed... Without using them they are lost, but for some cases my previous knowledge helped me to decide the method to follow. I believe there could be other ways, but personally I go for that.” (Berkin, Public Primary School)

“As I am an experienced teacher for 16 years, I use my old knowledge and what I lived in the class, I mean, my old practices. That way reminds me of what to do, what

to use as testing... Also, I try to remember which methods are more effective and my previous students liked..." (Cenk, Private Language School, Primary School)

"I generally think about what I did last year and previous years. Which methods were useful and not effective that much. Building on this, I believe is the easiest way and real. Otherwise, If I read something, that would not be suitable for my personality or teaching style and teaching philosophy. If I am more focused on process, for instance the methods that are more about the result will not suit my style and philosophy. Or another case may be choosing whether rubrics, portfolios or oral presentations are better for my context and learners and aims. Shortly, experience is a better source of learning and remembering that experience is the best and the most effective way to decide as far as I am concerned." (Doğa, Public Secondary School)

These teachers suggested that using prior knowledge and practice that they have is the way they opt for. Some of them signified that this way is more meaningful and efficient for them. While a few of them marked that they determine the method to the extent that they remember their testing course in the university, others stated that they choose based on their previous classroom practice as a tester.

Do you think that LTA training in undergraduate education is adequate for you as a language teacher?

Yes

No

If the answer is No, could you state in which areas of LTA do you think you need training?

During the interviews, when asked about whether training in LTA in university education is enough for them or not, 6 of them stated that the education was adequate for them. However, 14 of them noted that was not sufficient and when they were required to list the fields that they are in need of more training. Especially, the frequency of topics mentioned like statistics, using ELP and testing grammar is greater.

“When you ask this, I directly thought. Yes! I need absolutely. I never think that our training in 4 years of education is never enough for us. For none of us, actually. Of course, we can learn some things on the road and as we teach. But, testing is for instance specific. So, we need an expert to instruct it. Anyway, we need additional courses and one of them could be maybe statistics. I do not view myself as enough in this.” (Alican, Private Language School, Secondary School)

“I wish we received all areas of training in everything, but I know it is impossible to learn all of them. It is so vast and full of details. Example can be using more statistical things to evaluate our students. We only give scores and notes. That’s all. Also, for grammar. It is the major point of our teaching no matter how much we deny this. We only use traditional test formats: what is that?! Of course, test items and choices. It is more than that. We should be able to do that...” (Fatma, Public High School)

“Of course, not enough. Without doubt... I heard something new like ELP through social media, but I have no idea about that. Like how it is working and the ways to include it for assessing. Maybe, we could get training in such new developments.” (Eda, Public Secondary School)

“We could discover more about ELP. That is the thing I often come across nowadays online. In our books, there are some parts at the end of each unit like “can-do” sentences. Students check themselves and put ticks at the end of every topic. Such

types are, I guess, are examples, but I need to be more sure. It is really important for them to measure themselves. Of course, it is not that much objective to assess one's own, but it is better for awareness. Maybe we could take a short course on this." (Ayla, Public Secondary School)

As can be seen in the comments above, some teachers point that they should learn these topics that they see or hear either on social media or the internet. They address these topics as important to find out. On the other hand, other fields are mentioned like discovering more about testing listening, speaking and vocabulary.

"Nope. I mean that is not enough, of course. I wonder how it could be sufficient only in 1 course. Of course, many things could be added to an imaginary extra course curriculum like how to test our students' verbal skills or measuring their knowledge of words. We could be given more courses on many areas. These are the topics that I need more training." (Seher, Private Primary School)

"We know vocabulary plays critical role in language and we have to check that regularly. But, it does not necessarily have to be providing students with the word and asking to translate it into the mother language. Another thing almost all of us did was matching or completing the sentences. But now, everything changes, and we have to apply other ways like more motivating and engaging. We should be taught this. Well, I guess that is the first thing for me." (Alper, Private Language School, Secondary school)

"Today, we are all exposed to English as a dominant language and our students, too. When they watch movies, most of them say they watch it in English without much understanding. They always ask me how to develop their skills of listening. I say you have to listen to something at least for 30 minutes every day. Obviously, we

should do it in the class rather than only making them listen to audios for artificial purposes. Also, there should be ways to test their development in listening skill that is important for everyday life. We can be provided a course on this, I suppose.” (Merve, Private language school, Primary School)

“Unfortunately, nope of course. Based on my experience and intuition, we can be taught how to test our students in some skills like listening. I do not any other way rather than opening something and they fill in the blanks or mark the best choice. I want to learn more about that in a theory-based course.” (Okan, Private Primary School)

Like these 4 teachers, others articulated these topics due to the fact that they have limited knowledge and ways to test language skills and knowledge. As a result, they seek to get familiar with other ways of measuring learners differently.

Do you believe it is important for language teachers to be competent at LTA in order to increase the quality of learning and teaching?

When in-service EFL teachers were asked to state whether they regard LTA competence as essential for enhancing their learning and teaching practices, 13 of regarded it as highly important by justifying it through some reasons. While 4 of them said it is partially important, 3 of them stated they did not find it crucial at all.

“Definitely, yes! As teachers, we must be competent in almost all areas. When someone thinks of a competent teacher, she will think probably being good at teaching methods and techniques. To some extent, they will be enough, but teaching is not limited to that. It is beyond these features and one of them is testing without doubt. Also, these processes are all connected with each other. Without one, the other would be incomplete. What I am trying to say is this they are all related, and testing, learning and teaching are all parts of quality training. For me, after each

topic or unit I want to test whether my students have acquired the knowledge, or they have problems to handle these. If I had to put them in a pie chart, I would easily say that 40% of my time was on testing and the rest on teaching something. It took most of my time; that's why, we have to be qualified in LTA." (Gizay, Private language School, High School)

"I absolutely believe in the importance of testing as much as I believe in that of teaching. They are interrelated. For example, testing supports teaching, but if something is wrong with teaching, it can be detected through good testing, too. In a way, it serves a diagnosing tool for learning and teaching. Without testing, we cannot be sure of learning and teaching. I hope I could explain what I mean." (Okan, Private Primary School)

"Frankly speaking, I totally believe in this importance. No matter what our majors are as teachers, we have to be good at testing. We are all involved in preparing and applying these tests and education is comprehensive. When we talk about quality, it includes all things like testing and teaching, also learning. These three are all related. Although we were taught to focus on teaching methods and learning styles, they are not enough. We have to be holistic and should not forget testing aspect." (Berkin, Public Primary School)

As seen in these teachers' comments above, they point out the interconnectedness of teaching, testing and learning and they strongly believe that competency in LTA is important for them as teachers. Particularly, Gizay teacher made this connection through comparing the time allocated for test practices and teaching itself. In this way, she explained the relation between these processes. Also, Berkin teacher justified his belief through

holistic view of testing, teaching and learning processes. For Okan teacher, testing functions as a tool to detect and foster instruction and learning.

“Sure. Certainly. To increase the quality of work we deliver, we should be good at it. If we are not good at testing, how could we choose and use the best methods to assess? I think it is impossible. For example, if we do not know many methods, how could we use portfolios or checklist or maybe rubrics? Then, we would be using tests like fill in the gaps all the time. This would not be effective.” (Fatma, Public High School)

“I completely believe in the significance of assessment competency. We must be up to date in all fields. Technology improves every second, so everything like students and systems adapts to this and we should do the same. That includes being good at using technological devices at classes and using the most contemporary methods of testing rather than traditional ones as we were exposed to during our education as learners. We have to go beyond it.” (Doğa, Public Secondary School)

“I believe we should be competent, of course. If you ask me why I think in this way, please compare using old and ineffective methods of testing and current ones. They all test, but the newest ones test students in more engaging ways and capture them better. Being competent at testing involves this. The first thing ... comes into my mind is this reason.” (Merve, Private Language School, Primary School)

Like these three teachers, others also stated the importance and clarified themselves by means of employing proper ways of testing and assessing. Doğa teacher

explained that being updated is substantially important and one of them is assessment methods in that they should expose students to modern ways of testing. Others also mentioned this competency issue by making comparison between old and modern ways of testing like fill in the gaps and portfolios.

“One of tasks is to give reliable information about our learners and it is realized with good testing. You know, both teachers and schools have to give account of students’ success and failures. This is prevalent. If we cannot measure our students’ weak areas, how is it possible to improve these and give accurate information about them to our school managers and parents? Put it in a nutshell, all teachers should know this importance and be experts in tests. (Alper, Private Language School, Secondary School)

“Sure, I agree with this statement. Language teachers should be good in the field of LTA. Of course, for many grounds, it is essential. Firstly, it is integrative, I mean that teaching and testing connect to each other. Also, as language teachers, we have to report our students’ achievements and the problematic areas although it is demanding. Both determining these is demanding and trying to be sure of this information is difficult. We should be able to provide accurate information about students through accurate measurements and scoring ways. Shortly, testing expertise is a must for us.” (Seda, Private Secondary School)

These teachers believed that testing competence is essential on the grounds that they should provide different stakeholders of education with accurate information about learners’ strengths and weaknesses as well as overall achievement and failures. Therefore, they emphasized the cruciality of testing specialization.

As for others, 4 of language teachers stated they regard testing competence important, but to some extent. When they were asked the reason, they explained it through informing students.

“I believe it is important, but not that much. I mean moderately... I have the task of giving information about them perfectly. Which points are missing and why... (Alican, Private Language School, Secondary School)

“Let me think. I do not think that competency in LTA is not important. I mean it is important but not so important. We will use testing to announce to our students what they should study further also for ourselves to take notes what to do and plan in next teachings. I believe test is not only numbers but guides our students and us.”
(Melis, Private School, High School)

These teachers expressed they view competence in LTA as somewhat essential owing to the fact that they have the responsibility of keeping students advised of their learning and points to advance.

On the other hand, 3 of the in-service EFL teachers did not consider LTA competence as significant and approached testing as a separate thing.

“I do not think we have to be expert in testing. It is not our profession. We are expected to teach perfectly. Testing is another area and there are testers all over the world. I think it is not fair for us to be expected to be expert in this. It is the task of other professionals. Also, I do not think any direct connection between good teaching and testing.” (Selin, Private Language School, Primary School)

“For me, it is not important to increase education quality. I mean it is not our priority. For us, it is important to deliver the content on time and keep up with new methods

and ways to teach. I do not think that the quality of our teaching or learning is directly related to testing tools or results.” (Ada, Public High School)

“As far as I am concerned, we do not need to be expert in testing to increase the quality. Teaching, learning and testing are totally distinct from each other. Being expert in testing does not mean you are expert in teaching. I do not see any relation, honestly. (Mustafa, Private Primary School)

These participants did not hold the view that competence in language testing and assessment is required for making teaching and learning better. What was common in these 3 answers was that they did not relate the quality of these processes to the expertise in testing. Even, one of them made a sharp distinction between the two.

What kind of problems and difficulties you have experienced related to testing and assessment practices as a language teacher (e.g. preparation and administration of tests, students, reliability)? How do you overcome these problems?

In the interviews, in-service EFL teachers were first asked to express the difficulties they have had in terms of LTA practices. They articulated some areas of testing and assessment.

“For me, sometimes I have problems with preparing tests. It is both demanding you know and technical. We are not supposed to do this task. Which things I should be careful about in that there many things to think... You got my point, right?” (Dilan, Private Primary School)

“Let me think a bit. Probably, designing test from the beginning. Okay, I know the topics to include but about the distribution and types of questions, I find these difficult. Even now I do not know what to pay attention for while preparing it.” (Fatma, Public High School)

“The biggest problem is how to design and apply the test to students. It sounds easier, but I do not believe so.... Well, you have to include everything and every step carefully like principles, but I am not expert, so I sometimes choose ready-made tests, but they still need some arrangement. (Cenk, Private Language School, Primary School)

Especially, 10 of the interviewees stated the issue of “preparing tests” directly. They articulated the demanding nature of preparation of tests like concentrating on many aspects at one time.

“Testing is difficult on its own. The most striking one is administering the test. I mean the environment in which it takes...I want to give an example. When I choose some audio to test listening, there can be some problems with the connection or sound system unexpectedly. It is the smallest problem that I found now. In a real class, many things happen during the administration.” (Sezen, Public Primary School)

“Technically, we prepare, print the tests and apply them. Until that moment, everything is okay. But we can’t control everything like our students. Some students may have obvious problems on this day, and I have dilemma whether to delay the test or not. If it was a quiz, that would be okay maybe to delay it. But if it is like end of term exam or serious one, what could I do? I can’t control the motivation or problems of each and every student. There are many possibilities I mean maybe they have problems with their friends or sleepy and tired. I am sure they will affect

their scores, but I have real difficulties to choose the best decision at those moments.” (Ayla, Public Secondary School)

“You know teaching profession is full of unexpected things because of students and classroom environment. Though the test looks perfect in fulfilling instructional aims, sometimes it may not achieve this. Some of the students may look too tired to focus on the exam or look anxious while few of them are ill. Under these conditions, it makes my job more challenging.” (Sezen, Public Primary School)

“It is really difficult to ensure reliability and validity issues. You know it is so hard to obtain similar results if you used the same test to the same students at different times. I cannot control all the things in the exam. Similarly, covering all the topics and the proficiency we wanted to measure... I find them difficult.” (Eda, Public Secondary School)

“I suppose it is reliability. I talk this to my colleagues, and we discuss this for a long time. We have problems in providing reliability. Think about this: many factors affect the test results like students’ motivations, sound system, noise and the quality of photocopy...” (Seher, Private Primary School)

Other teachers addressed the problems of applying tests in the real classroom by referring to some aspects of reliability issues, too. Most of them stated that they cannot go over everything in the classroom, which is students themselves and environment like computer, sound system and photocopy. They experience hesitations about what to do with the exam or exam results. Also, some gave voice to the problem of validity including some types like content and construct without directly telling these terms.

“Well, we use tests and obtain results as usual, but I think the most difficult thing is to explain this to students, that is, feedback. It requires skills I believe. We should be motivating and telling the exact points to them in correct ways. I have problems with this. I need to think much time: what to say, how to say this... so that they will not be discouraged or unhappy.” (Ahmet, Public Primary School)

“In my experience of 10 years, I have experienced many problems in testing like giving feedback, reliability and of course using the tests at the classroom. It is natural I believe as it has many dimensions to be careful. As an instance, students are so sensitive, and we have to give correct feedback not to lose them or their motivation. Certain students may take this as personally. However, we have to be realistic at the same time while telling their mistakes or errors.” (Alper, Private Language School, Secondary School)

As seen above, few of them mention the challenge of feedback on the grounds that they consider the motivation and well-being of learners. Particularly, Alper teacher stated the risk of learners' taking these comments personally and to avoid such situations, they have to think more on this issue.

Furthermore, they were asked to express the ways dealing with these problems. What most of them stated was consulting to other workmates, conducting research as well as reading relevant things online. Three of them stated that they ignored the problems if they were not critical.

“... Okay... I got the point. Generally, I ask for some advice from my colleagues for possible ways like what they did how they gave feedback...Of course, it changes in every class but to solve the problem I need other perceptions. (Ahmet, Public Primary School)

“... I talk this to my colleagues, and we discuss this for a long time... when we talked about these common things, first I was relieved that I was not only one who lived this. Then, we try to find alternative ways to the problems.” (Seher, Private Primary School)

“... Sometimes, I consult to other English teachers because they know the context and students, too. If I talked to someone else even though he could be an expert, he would not know the realities like students, background, parents, and administration. So, it is more valuable and useful according to me.” (Ada, Public High School)

“... I like collaborating with other teachers because we know each other, and our preferences and dislikes. They know my strengths and I know theirs, too as we are working together for long years. It makes more sense, I mean..., a greater sense of accountability. In this way, I feel more confident about the problems. To make it clearer, recently I had a problem in one of my classes about the quiz I prepared. Some students objected to it, and they said it was not fair to be measured without announcement. I could not make it and I discussed this with my workmates about the solutions how to decide such quizzes without having conflict with students...” (Mustafa, Private Primary School)

As seen above, some teachers preferred sharing the problems with their workmates for various reasons like their awareness of the context, students, school, feeling more secure and confident, and knowledge of strengths and preferences.

“... When I faced a problem in my class, I directly “google” it and try to read some newsgroup and find similar cases and solutions. This way makes me feel relaxed

and secure. If I shared it with another familiar teacher, I would be embarrassed. Reading others' comments and experience is effective to overcome those." (Selin, Private Language School, Primary School)

"... I guess I try to find and read similar problems in online platforms like teachers' groups. Reading these similar problems and solutions provides me different perceptions and new ways of preparing and using tests." (Sezen, Public Primary School)

"... I am an experienced teacher and what I have learned from these is that you should not talk such problems to the other teachers at school. I know it may sound awkward, but it is the reality we must face... Anyway, of course I like sharing something with other people, but I believe in the importance of distance among colleagues. Anyway, I like searching on the net and reading forums." (Alper, Private Language School, Secondary School)

These teachers revealed that they choose to do research and read relevant forums or newsgroup to learn possible ways to handle the problems. Particularly, Alper teacher said that he was strongly opposed to the idea of asking colleagues based on his prior experience.

"If I had some problems including testing and assessment, I would just have some time to think and try to relax myself. As part of our jobs, we live lots of problems and I just ignore the problem and take it easy to solve it. After some time passes, they disappear." (Berkin, Public Primary School)

“... I experience so many problems that I try to say: Be calm and let the problems flow. They are natural. Sometimes the level of problems change, but they are not so critical.” (Merve, Private Language School, Primary School)

“We have many problems in our classrooms like in our materials or tests or just during teaching. Of course, there will be problems and if it is not serious or continuous, I think I should not exaggerate it. However, if it always occurs, I would think of serious solutions.” (Doğa, Public Secondary School)

These three teachers explained that problems are casual parts of their occupations and if they were not so critical, they would ignore it for some time to solve it. However, Doğa teacher stated that if the same problems happen many times, she will try to come up with essential solutions.

Do you have any questions or comments about the subject?

The interviewees were directed the question of whether they have any further remarks or inquiries and only 2 of them asked questions.

“I have a prompt question to you: When I think of these aspects that we talked in the questions and the survey I did, I wanted to improve myself more in language testing knowledge. How can I develop myself? Could you suggest me the best thing to do this?” (Dilan, Private Primary School)

“Thank you for giving that opportunity to me, firstly. You increased my consciousness and I wonder in which ways I could advance my knowledge in testing. It has many sub-topics, and I am interested in learning many of them as much as

possible to apply in my classes. Of course, I can read some books, but I want a more interactive and practical way of doing so.” (Seher, Private Primary School)

Based on these questions, these two teachers were interested in discovering more about LTA and they inquired these ways. They seek to increase their knowledge in this field and one of them pronounced that her awareness has increased.

Discussion

The results of the present study are discussed in relation to relevant literature. The chapter is organized according to the research questions of the study. There were four research questions in the present study as stated in the preceding chapter. After each of the research questions is given, a summary of the findings is provided. It is followed by relevant studies for comparing and contrasting the findings. Upon discussing the findings, suggestions for an EFL in-service teacher training programs are made.

Discussion of the 1st Research Question

In-service EFL Teachers' Evaluation of Undergraduate Training in LTA

The first research question of the study was “How do in-service EFL teachers evaluate their undergraduate training in language testing and assessment? Do they perceive this training as adequate?” There were four components to be examined under this question: (1) Classroom-focused LTA, (2) Knowledge of Testing and Assessment, (3) Purposes of Testing and (4) Content and Concepts of LTA. Data gathered through the questionnaire was analyzed so as to answer this question.

To start with the first sub-category that is Classroom-focused LTA, the results showed that they find the training they received in the university as inadequate with the mean score of 1.39. The top three items the EFL teachers stated a higher amount of training were “using European Language Portfolio”, “stages of language test construction” and “preparing proficiency tests”. It was found that nearly 40% of them evaluated the training as

satisfying in these items. On the other hand, “interpreting test scores” has been found to have the lowest mean score as 19% of them revealed no training in this item. “Preparing diagnostic” and “preparing achievement tests” had the second lowest mean score with 1.26 and nearly 17% of them stated they were not given training. Another item that has been found to be the one of the lowest scores was “using self / peer assessment” and nearly 19% mentioned no training at all. Another aspect was “giving feedback to students based on information from tests/assessment” and 14% of the respondents revealed no training at all. All in all, in accordance with the results for 18 items in this sub-category, in-service English language teachers regard the training received not sufficient. A similar pattern of results was obtained in the other studies in the literature (Hasselgreen, Carlsen, & Helness, 2004; Vogt & Tsagari, 2014). In Vogt and Tsagari’s (2014) study, teachers’ training was found to be inadequate and underdeveloped in this domain since 33.7% and 34.6% of them reported no training and a little training for the given domain overall, respectively. Moreover, they pointed out how underdeveloped English teachers’ LAL was on the grounds that receiving training in LTA plays a role in the level and development of LAL.

The overall finding of the current study and the studies of Hasselgreen, Carlsen, and Helness (2004) and Vogt and Tsagari (2014) was found to be consistent. More specifically, in Hasselgreen, Carlsen, and Helness’ (2004) study, some items like using ready-made tests, giving feedback and using informal, continuous assessment were found to have been carried out without training. As to using portfolios (here defined as ELP), it was found to be as an untrained area. Similarly, in the study of and Vogt and Tsagari (2014), alternative forms of assessment, such as the “ELP or Portfolio” were highlighted as the areas in which either no or little training was reported. Also, Ballidag and Inan Karagul’s (2021) study revealed that around a third of the participants did not receive any type of training in ELP. However, in the current study, ELP was found to have the highest amount of training compared to other items. Considering the importance of the European Language Portfolio for reinforcing language learning skills as well as the growth of learner autonomy,

plurilingualism and intercultural awareness, training level in this component is highly critical. Another critical point is self and peer assessment in which little training was reported by in-service EFL teachers. These two types of alternatives in assessment that are self and peer assessment are crucial for teachers to utilize given that they serve a function in autonomy and motivation as well as making decisions for one's own or others' learning and work. Provided that teachers employ these alternatives in assessment in their classroom, they enable students to take more responsibility, work collaboratively and engage more actively during the education process. Consequently, getting training in these areas including the importance of them and ways to employ them are noteworthy for teachers.

As for the interpretation of test scores that is also important for teachers to include in their assessment processes, teachers have to accomplish it after conducting any type of assessment owing to the fact that they have to grade and make important decisions like placement, allocating more time to the content and improving areas that need more attention. Just obtaining results from any type of assessment is not adequate and does not ensure that the process is completed. Moreover, in-service EFL teachers are expected to deliver feedback in accordance with the information elicited from assessment effectively. This includes explaining how the grade is calculated, identifying and rewarding strong aspects, guiding and motivating students to improve their learning perpetually as well as ability to evaluate themselves to check where they are in the learning process. Considering the lack of adequate level of training in these aspects and the substantial role of them for both teachers' profession and learners' development, these should be incorporated into LTA courses more.

The finding that training in some types of tests like diagnostic and achievement was not satisfactory was significant due to the fact that the former fosters the strengths and weaknesses of learners by determining certain areas and the latter gives information about learning and teaching to alter something if required. Both of them have formative roles and enhance learning and teaching continuum in addition to making them take actions during

the process. Lack of training in these may have negative washback on learning and teaching processes and it could be explained by the fact that there is a tendency to include summative assessment types more as there is an exam-orientation in Turkey.

Additionally, Salami and Alharthi (2022) found that the overall training level in this category was not enough for LAL development, which accords with the finding of this section of the study. However, the level of training received in “using self or peer assessment” was higher than that of the current study. This might be owing to the differences between samples of these two studies.

The second sub-category is “Knowledge of Testing and Assessment” and the findings revealed that the training they received in undergraduate education was nearly sufficient with the mean score of 1.56. When the training level found for this category and the previous one is compared, it is found that there is a slight difference in the training levels between “Knowledge of Testing and Assessment” and “Classroom-focused LTA”. Considering the effect of the level of training on the level of LAL, their LAL could be stated as not well-developed, but slightly developed. The item with the highest mean score was “objective / subjective testing” and around half of the participants enunciated the level to be sufficient. “Computer-based testing” and “norm / criterion-referenced assessment” are the other items for which around 40% of them reported adequate level of training. On the contrary, for “informal / formal assessment” that had the lowest mean score approximately a fifth of the participants pointed out no training at all. In the same vein, “discrete point / integrative testing” was the other item with low mean score as nearly a fifth of them expressed no training. When the importance of these items for both test-takers and educators is considered, it can be stated that for giving a better understanding of student learning and achievement, all types of assessment should be incorporated into the training offered to teachers and teacher candidates.

The other sub-component is “Purposes of Testing” and the training level with the mean score of 1.51 was revealed to be slightly adequate in this category. For all 7 items,

the mean scores were found to be closer to one another. It could be argued that LAL level of them regarding this field is not well-developed. Almost four out of ten participants highlighted sufficient level of training in “placing students onto programs, courses, etc.”. In a similar manner, other areas in which sufficient level of training was reported were “measuring general ability to learn a foreign language” and “awarding final certificates”. It was really interesting to find that “awarding final certificates” was paid attention in ELTE courses though the learners are not assigned certificates in the Ministry of Education. The items which have slightly lower mean scores than others were “giving grades” and “identifying what has been learned”. The overall level was found to be slightly higher than that of previously reported level (Vogt & Tsagari, 2014). In their study, there were four items under this category and in most of them respondents stated that they did not receive any training in them. More specifically, almost 40% of them stated no training was offered to them in these items. Especially, for “giving grades” and “placing students onto courses, programs” around half of them mentioned any type of training was not given to them. Also, “awarding final certificates” was another item in which nearly 60% of them similarly stated no training. Considering the findings, it can be stated that for some items there are differences in the findings of their study and current study. Whereas in the current study around 40% of in-service EFL teachers stated adequate level of training in “placing students onto programs, courses, etc.”, in their study half of them stated no training in this item. However, for other items like “giving grades”, training was not found to be enough. Based on their study, Vogt and Tsagari stated that principal constituents of LAL of teachers are not enough, and they put forward that these constituents are likely to be learned in teaching profession.

In the study conducted by Salami and Alharthi (2022), EFL instructors' perceived level of LAL was investigated including their training levels in different domains. It was found that on the domain of “Purposes of Testing”, EFL teachers stated they had received the minimum amount of training in “placing students into courses, programs”, and “awarding

final certificates". In another study carried out by Ballidag and Inan Karagul (2021), similarly "awarding final certificates" was the area in which most of the participants revealed either no or little training. Therefore, the areas that seemed to be less developed than others differ in their studies and the present study.

When it comes to the last sub-category which is "Content and Concepts of LTA", it is found out that the training they received in their undergraduate education was almost enough with the mean score of 1.73. The items they reported a higher amount of training were "authenticity", "using statistics to study the quality of tests / assessment", "practicality", "different test items/question types to test pronunciation in English" "washback", "different test items/task types to test reading in English" and "validity". In these items, most of them signified either advanced or sufficient training. Receiving adequate level of training in some of the principles like validity, practicality and authenticity is vital for the development of LTA knowledge and assessment literacy. Considering why validity is so important for assessment on the grounds that it gauges what it seeks to gauge, being trained in this principle is necessary. However, reliability is also necessary along with validity and other principles, and in the findings, it could be seen that the level of training for them can be also accepted as enough. Drawing accurate conclusions from any type of assessment depends on these principles and being provided enough training in them is essential.

Conversely, "testing speaking in English" was found to have the lowest mean score due to the fact that 18% of them revealed no training in this item. Speaking skill is of utmost importance for language learners and assessment of this productive skill is also of paramount importance for language teachers. It is not just giving students some questions and asking them to speak but requires many techniques and skills. Having been trained in this aspect allows different types of tests and assessments to be utilized as well as designing and applying valid and reliable tests. Unless teachers are trained in this to much extent, it is highly possible that sound assessment practices could not be conducted and

limited test types of speaking could be included rather than using different types like interactive, intensive and extensive, which could hinder the effective education practices. Without satisfactory training, there could be challenges in the construction, design, application of assessment. It could be stated that these practices are carried out without much training.

Compared to the prior three sub-categories, participants reported that they had received the greatest amount of training in this field, which is in accord with the finding of the work of Salami and Alharthi (2022). As for the study of Ballidag and Inan Karagul (2021), it was revealed that the training was found to be a little (1-2 days) for all items. The item in which most training was expressed was “testing and assessing microlinguistic aspects (grammar/vocabulary)”, which differed from the present study’s finding. Another difference is that the level of training that is mentioned for productive and receptive skills in their study is the same. However, the present study revealed that testing receptive skills have higher mean scores than productive skills, especially speaking skill. This is consistent with the work of Hasselgreen et al. (2004) in that training in productive skills is stated to be low and therefore they put forward the idea that these assessments are likely to be conducted without training. Similarly, “using statistics”, “assessing culture”, “reliability” and “assessing integrated skills” were also mentioned to be carried out without taking training. Also, in their study there are differences in the amount of training expressed for different items of this domain, which means the training provided to them is less in certain items despite comparably more in others.

In Vogt and Tsagari’s (2014) study, 29.4% and 35.1% of them reported no training and a little training for the given domain overall, respectively. Particularly, almost 60% of them expressed that no training was provided to them in “using statistics”. Therefore, it was highlighted to be one of the critical neglected areas. Furthermore, more than one third of the participants stated they did not get any training in “reliability” and “validity. Accordingly, they suggested their LAL was not developed well since training level which was found

inadequate shapes their LAL level. This differs from the finding presented here in that the level of training was almost satisfactory.

All in all, the first research question of the current study sought to analyze in-service EFL teachers' evaluation of their undergraduate training in language testing and assessment and the results of all 4 subparts showed that the overall mean score was 1.57. Accordingly, this reveals that overall in-service EFL teachers evaluate their undergraduate training in LTA almost satisfactory, though not completely. This suggests that the level of training was slightly sufficient and considering this level, it could be stated that their LAL is developed to some extent, though not well-developed. In other words, they can be stated to have a certain level of LAL despite lacking a very-advanced level of knowledge of LTA. This is in agreement with that obtained in the study of Vogt and Tsagari (2014) who also reported not well-developed level of LTA literacy among teachers across Europe. Also, this corroborates the overall findings of Ballidag and Inan Karagul (2021) and Salami and Alharthi (2022) in that LAL level was found to be not developed well based on not taking advanced or sufficient training in LTA components. In their study, most respondents stated either no or little training in these components and the overall training level found in the present study was slightly higher than studies of Ballidag and Inan Karagul (2021) and Salami and Alharthi (2022).

As for the study of Hasselgreen et al. (2004), it did not yield similar results since they found that in three areas of LTA teachers did not get much training and conduct their assessments without much previous formal training. Also, work of Kaya and Mede (2021) revealed that instructors in Turkey were found to have a higher level of LTA knowledge; however, whether receiving training or not in LTA did not affect their knowledge level. Therefore, the findings of the present study differ from these two studies.

More specifically, in the current study, in-service EFL teachers consider having received enough level of training in different components of LTA; however, they regard the amount of their LAL training mostly either none or little for some aspects like "interpreting

test scores”, “preparing diagnostic”, “preparing achievement tests”, “using self / peer assessment”, “informal / formal assessment”, “discrete point / integrative testing” and “testing speaking”.

In the present study, the greatest amount of training compared to other three sub-domains was found in “content and concepts of LTA”, and the least amount of training was reported for “classroom-focused LTA”. In the study of Vogt and Tzagari (2014), the most developed area is “content and concepts of LTA”, too. However, the least developed in their study is the domain of “purposes of testing” in which almost 40% of them reported to have been offered no training at all. Thus, crucial aspects of LTA literacy including “giving grades”, “placing students” and “awarding certificates” are expressed not to be developed. As Popham (2009) noted, to enhance the LTA knowledge and LAL levels of teachers, providing enough training in undergraduate education is critical.

Discussion of the 2nd Research Question

In-service EFL Teachers’ Needs for In-service Training in LTA

The second research question of this study was “To what extent do in-service EFL teachers perceive a need for in-service training in language testing and assessment?”. To answer this question, four aspects were examined: (1) Classroom-focused LTA, (2) Knowledge of Testing and Assessment, (3) Purposes of Testing, and (4) Content and Concepts of LTA. The data collected through the questionnaire was analyzed to provide an answer.

The first domain is Classroom-focused LTA and the results revealed that they need further basic in-service training in this field with a mean score of 1.43. The results showed that the top areas the EFL teachers wanted to improve themselves were “preparing achievement tests”, “preparing placement tests”, “preparing proficiency tests”, “preparing progress tests”, “preparing diagnostic tests” and “using self/peer assessment”. For these items, nearly 80% of them stated further moderate and basic level training needs with similar

percentages. As for the item of “using European Language Portfolio”, participants enunciated the need for advanced and moderate level of in-service training with around 11% and 40%, respectively. Contrariwise, “grading” has been found to have the lowest mean score and “adapting ready-made tests for the needs of students” is found to have the second lowest mean score on the grounds that nearly one fifth of the respondents did not request any further training in the given items. Another item was “scoring” due to the fact that nearly 20% of them did not mention any need for further in-service training. Therefore, according to the findings for 18 items in this domain, in-service English language teachers’ need for further in-service training is found to be at a low level. When the finding of further training needs of them in this domain and that of received training in the same domain is compared, it is interesting to find out that they only express a basic need for further training in this area even though they generally connotate the training received in pre-service education to be inadequate. One possible reason for this rather contradictory outcome could be that they are not fully aware of the significance of LTA for their current careers and language teaching practices. Additionally, these teachers may not have the time or resources to pursue further training in this area. Furthermore, there could be other relevant factors which influence why these language teachers may not seek further training in LTA including lack of institutional support like access to professional development activities or relevant conferences or training workshops, not seeing the immediate benefits of further training in LTA, overreliance of standardized tests like using ready-made proficiency tests. As the level of training that they received affects the level of LAL, their LAL could be described as not well-developed and not wishing for extra training in this domain could also affect the level of LAL negatively.

However, there were some items in which the need for further training need was expressed in parallel with the insufficient training received. These items were “preparing diagnostic tests”, “preparing achievement tests” and “using self/peer assessment”. Therefore, it could be stated that these teachers are aware of the significance of them for

themselves for many reasons like identifying learners' strengths and weaknesses, assessing students' progress, achievement and encouraging student involvement in the assessment process and development of metacognitive skills. These, in turn, lead to more effective learning outcomes by providing teachers with valuable insights into students' learning processes and outcomes, evaluating the effectiveness of their teaching methods, tailoring their teaching to meet the needs of learners, and making adjustments as needed.

On the other hand, there were some items for which further training need expressed despite the fact that the training that was stated to have been received was mentioned to be sufficient. Specifically, for the item of "using European Language Portfolio", though nearly 40% of them regarded the training received as sufficient, 40% of them expressed moderate level of further in-service training. It is possible that these participants are aware of the importance of ELP along with its use and benefits in that it will enable them to assess their students' language skills more accurately and improve communication with other language teachers as it is a standardized tool that is used across Europe as well as supporting students to set realistic language learning goals by enabling them to take active roles on their own learning like setting goals and reflecting upon their own progress. Their further training demand in this item could be explained considering its being a comprehensive framework for the assessment of language proficiency and its role as a communicative tool with other colleagues all over the Europe in addition to enabling them to set achievable learning goals. Similarly, "preparing proficiency tests" is another area in which almost 40% of the participants revealed their moderate level of further training need although almost the same percentage of them found the training that they received adequate. A possible explanation for this might be that teachers have a sense of importance of this type of test since receiving more training in this aspect will provide them with designing more accurate and reliable tests that offer a clear indication of their students' language proficiency. Also, it will facilitate the understanding of teachers how to design,

administer, and evaluate proficiency tests with the aim of meeting students' needs and tracking their progress over time.

This study supports results from previous studies (Hasselgreen, Carlsen, & Helness, 2004; Vogt & Tsagari, 2014) in that there is a need for further training in the given domain. In Vogt and Tsagari's (2014) study, need for in-service teacher training in this domain of Language Testing and Assessment (LTA) across the board was found; however, specific content areas and extent of training required are said to vary depending on the assessment requirements of different educational contexts. More specifically, in their study 36.5% of them highlighted their wish to receive advanced training for this domain and 28.3% stated the need for a little training in almost all items included in this domain. Nevertheless, training need for some items is expressed as more immediate than the others when the findings of the current study and their study were compared. In their study, almost half of the teachers (47.7%) expressed a need for more advanced training in "preparing classroom tests" despite around one-third (33.8%) having already received advanced training. Nonetheless, for "ready-made tests" almost half of them did not want to get further training though nearly 40% of them stated to have received no training at all. As for "giving feedback to students based on information from tests or assessment", most of them stated to have received certain level or advanced training, but they still wished to receive more training with 41% for advanced training need. Furthermore, in areas like "self- and peer-assessment", "informal assessment" and "using portfolios such as the European Language Portfolio for assessment", need for advanced training was pointed out (41.4%, 44.7% and 48.5%, respectively). For these alternatives in assessment, the study of Hasselgreen et al. (2004) yielded similar results in terms of prioritizing them for future training needs. As stated by Vogt and Tsagari (2014), this could stem from the fact that this type of assessment is regarded as a new area of LTA by language teachers; therefore, training need is uttered clearly. In the same vein, in the present study, "using self/peer assessment" and "using European Language Portfolio" are among the areas that teachers wish to develop

themselves more by having moderate level of further training. In this regard, the extent to which they seek further training in these two items differs slightly between the study of Vogt and Tsagari's (2014) and the current study. All in all, despite the need for further training in the given domain, there are different priorities for different items between the two studies and the extent of the training need found in the study of Vogt and Tsagari (2014) was slightly higher than that of the current study.

In the study conducted by Salami and Alharthi (2022), it was found out that the overall training need in this sub-domain was more than the other two in their questionnaire. The area that they also wished to receive the most training was "using self- or peer assessment". The second area for which training need is expressed more is "preparing classroom tests". Also, similar to the study of Vogt and Tsagari (2014), there was a desire to get basic and advanced training in "giving feedback to students based on information from assessments" in spite of the fact that most of the participants revealed to have training. These results seem to be consistent with the study of Ballidag and Inan Karagul's (2021) who found that "using the European Language Portfolio" and "using self-or peer-assessment" are the reported fields for extra training. Also, the lowest need for future training was found for "using ready-made tests". In consistent with Vogt and Tsagari's (2014) study, "using ready-made tests" is not among the areas that are aimed to be developed more. It could be related to their inclination to trust available ready-made tests, which could account for its being the least wished area for training need.

As to the second domain, which is Knowledge of Testing and Assessment, the results showed that they need moderate level of in-service training in this field with the mean score of 1.65. EFL language teachers expressed the wish to develop themselves more in some components of the domain. The top components were "alternative assessment", "formative / summative assessment" and "informal / formal assessment" and 36.3%, 47.3% and 50.7% of them indicated moderate level of extra training need, respectively for these items. For other items which are "objective / subjective testing" and "norm / criterion-

referenced assessment”, most of them also reported need for moderate to advanced level of extra training. Contrarily, “approaches to language testing” was found to be the least desired area for further training need and it was followed by “direct / indirect testing” and “computer-based testing”. Around 10% of the participants did not report any kind of extra training need for these three items. Accordingly, for 9 items in this domain, need for extra in-service training is reported to be at intermediate level. When this outcome was evaluated considering the finding of training level received in the same domain, it was found that though the training offered to them was found almost sufficient, they still wish for extra training at a moderate level. Though these EFL teachers might feel that they have received almost adequate level of training in that domain, there could be some reasons why they wish additional in-service training like the evolving nature of testing and assessment field and keeping up with this trend, addressing different learner needs and using assessment tools more effectively. Particularly, it could be attributed to the fact that the field of LTA is constantly evolving, and it is critical for teachers to follow the latest developments in the field to ensure that both themselves and learners are provided with the most recent effective and accurate assessments and best practices. Also, considering the variety in the learner population and their needs, in-service training can equip them with a better understanding of assessment principles and techniques besides teaching them how to use these efficiently. Lastly, it is possible that they are aware of the fact that their LAL is developed to some extent but not completely, which could account for their wish to be offered extra moderate level training.

A comparison of the training need expressed for this domain and that of the previous one showed that there is a slight difference in the levels in that participants expressed slightly more training need in “Knowledge of Testing and Assessment” field. In this domain, there were certain aspects where the inadequacy of the received training was pointed out and a requirement for additional training was expressed. One of them was “informal / formal assessment” in which almost 20% of them reported to have got no training at all and half of

them revealed apparent intermediate need for extra training and more than 10% of them expressed advanced level of extra training need. On the other hand, even though the level of training that was received in some items was found to be sufficient, there was a need for extra training. "Objective / subjective testing" is one of the areas in which sufficient amount of training was given to them, but still moderate to advanced training need was enunciated. It may be explained by the fact that differentiating between two is substantially important for ensuring reliable, fair and valid assessments along with providing alignment with instructional goals and objectives. Incorporating these into the training offered to teacher candidates and teachers is crucial taking these factors into account.

The other domain is "Purposes of Testing" and the findings showed that in-service EFL teachers need intermediate level of extra training in this domain with the mean score of 1.86. It was found that the main areas that they need to develop themselves more were "identifying what has been learned", "measuring general ability to learn a foreign language", and "awarding final certificates". For these items, the majority of the participants, which is nearly 75 % of them, wished to develop their knowledge more by asking for moderate to advanced level of training. Moreover, for the other aspects including "testing competence in a language", "placing students onto programs, courses, etc." and "finding out what needs to be learned/taught", around 60% of the participants regarded their training need moderate to advanced level. On the contrary, "giving grades" was the least desired aspect for further training as nearly one fifth of them did not wish for any additional training. Consequently, for the 7 items of this domain, these EFL teachers pointed out the need for extra training at the intermediate level.

"Identifying what has been learned" was one of the aspects in which additional training was necessary as expressed alongside inadequate training already received. On the other hand, there were some items for which further training need was not expressed though the training that was stated to have been received was mentioned to be insufficient. For instance, "giving grades" was the item in which nearly one fifth of the participants did

not report any training offered to them. Surprisingly, almost the same percentage of them did not wish for any further training in this item. There were other items in which extra training need was reported even though the training offered to them was stated to be sufficient like “placing students onto programs, courses, etc.” and “awarding final certificates”. It is especially interesting that for “awarding final certificates”, which was the area in which sufficient training was stated to be taken, they wished to receive more training intermediate to advanced level. This is an unanticipated finding in some ways. First of all, the learners are not assigned certificates in the Ministry of Education; therefore, finding out the fact that the training given to them in this area was sufficient was surprising. Secondly, even though this is not applied by language teachers, they seek to receive more training. It was really interesting to find that “awarding final certificates” was paid attention in ELTE courses though the learners are not assigned certificates in the Ministry of Education. This could imply that maybe these participants give importance to final certificates as these can be used as a motivation for students to maintain their learning and as a credential for future academic and professional opportunities. All in all, if these participants’ level of training that they received in their undergraduate education, which is slightly sufficient with the mean score of 1.51, and their need for extra training in the same field with the mean score of 1.86 are considered, it could be seen that they are conscious of the necessity of knowledge of different aims of testing for different types of tests, students and educational contexts. They could have realized that their training and LAL level in this domain is not well-developed, but rather somewhat developed. It could also explain their desire to get extra training to enhance their level of LAL. Also, maybe they can be aware of the fact that just being knowledgeable about different test types and tools of assessment does not provide sound assessment practices. Rather, it is vital to be informed about diverse purposes of assessment in that whether the test aims to reveal what needs to be taught or place learners into appropriate levels of courses or programs. Even though there are reliable and valid forms of tests that they can employ, it does not guarantee that they can yield valid and reliable outcomes if they do not consider or conform to the purposes of conducting the test.

Thus, they may feel the need to take extra training to help them make informed decisions and improve the quality of their educational practices.

This is also in accord with the previous study conducted by Vogt and Tsagari (2014) in terms of the finding revealing teachers' further need for getting training in the domain of "Purposes of Testing". However, the level of training need for different items of this domain differs between two studies. In their study, it was found that on average 60% of them asked for a little to advanced level of further training in this domain and as for the training that is reported to be received, it was found to be the least advanced area since on average 40% of them stated not to have received any type of training in this field. Especially, for "placing students onto courses, programs" (50.5%), and "awarding final certificates", most of the participants (almost 70%) wish for basic to advanced level of further training. As for "giving grades", a third of them did not desire any extra training related to it. This is consistent with the finding of the current study in which the least desired item for future training was found for this item and nearly 20% of them did not enunciate any extra training need. Similarly, the study of Ballidag and Inan Karagul (2021) also revealed that for "giving grades", participants did not even wish for basic training. It was found to be intriguing by the researchers as it did not appear to be a crucial area of teacher education despite its being a fundamental task for teachers. Hence, they suggested that items like "giving grades" and "awarding final certificates" are likely to be underdeveloped and these could be learned during their teaching experience. As they stated, it would be compelling to examine the learning of that on the job like the way teachers learn "giving grades" on their occupation when taking the reluctance of them for training in this item into account.

The study carried out by Salami and Alharthi (2022) showed that there was a training need in this domain, but they do not demand getting much training for this domain. Compared to the other items, the item for which much training need is emphasized more is "awarding students final certificates". This corroborates the finding of the present study where more than 70% of the participants highlighted intermediate to advanced level of

training need to develop their LAL levels regarding this. "Placing students into courses, programs, etc." also follows this with the second highest mean score for further training and most of them demanded training at the low level. This training demand was found to be less than that of in the present study.

As for the study of Hasselgreen et. al (2004) and Ballidag and Inan Karagul (2021), it is found that teachers need to have training in general including the domain of "Purposes of Testing"; however, it is not found to be "urgent"; in other words, the need for further training was not at advanced level and not among the priorities. Hence, the result of the current study for this domain matches the finding of their study overall; however, the level of wish for extra training in this study is found to be higher and more "immediate" than what was found in their studies. Especially, in the study of Hasselgreen et. al (2004), almost 40% of them were found to have stated that they had reservations about getting extra training. Also, they were found to conduct "placing students onto courses, programs" and "awarding certificates" without proper training, which was more prevalent than conducting tests for "diagnostic" or "giving grades" purposes. Among the items of the domain, for "finding out what needs to be taught" the expressed need is slightly more than the others. This is totally in line with Ballidag and Inan Karagul's (2021) study.

Regarding the last domain that is "Content and Concepts of LTA", the study has demonstrated that these in-service EFL teachers regard their future training level as moderate level with the mean score of 1.94. The results demonstrated that the top four areas the EFL teachers wanted to improve themselves were "testing integrated language skills", "practicality", "reliability" and "validity". For these areas, nearly 75% of them demanded further moderate and advanced level of training. As for the item of "testing pronunciation in English", most respondents revealed the need for advanced or intermediate level of training. The other items like "washback" and "authenticity" were also highlighted for future training as around 40% of them reported need for training at advanced

level. As for “testing listening in English”, “testing vocabulary in English” and “different test items/task types to test grammar in English”, advanced level of training was also demanded by around 30% of the participants. Contrarily, the least desired aspects for extra training were “different test items/task types to test speaking in English” and “different test items/task types to test writing in English” in that around 10% of them did not want to be offered any extra training on these aspects. All in all, when the result for 22 items of this domain was evaluated, it was seen that in-service EFL teachers were in need of intermediate level of extra training. Upon examination of the training level attained in the same domain, it was discovered that in spite of receiving almost adequate training, the individuals desired more moderate-level training. Despite feeling almost satisfied with the training they received, the EFL teachers may have valid reasons for wanting more in-service training. For instance, they may wish to deepen their understanding of the subject, refresh their knowledge, learn new strategies and techniques, keep up with the latest developments and gain more confidence in their abilities to utilize them in their professions. Moreover, they might consider that the initial training that was offered to them was adequate to some extent, but that is not addressing their specific needs or responding to concerns related to their own teaching contexts sufficiently. Last but not least, they are likely to consider that they have somewhat advanced level of LAL for this area, but not fully advanced, which could be attributed to their desire for more training.

For some items, extra training need was demanded despite almost sufficient level of training offered to them. For “practicality”, even though nearly 40% of them perceived the received training as sufficient, nearly the same percentage of them expressed advanced level of further in-service training. It could stem from the fact that there is more likelihood for practical assessments to be utilized regularly in the classroom, which provides language teachers with regular and reliable information about learners' language abilities. Also, as practical assessments are easy to apply and score, they reduce the burden on teachers and therefore let them concentrate more on teaching and giving feedback to students

instead of allocating most of their time to assessment application and scoring. Taking these into consideration, these teachers might have asked for more education in this concept. As for “validity”, despite the fact that almost 60% of them stated either moderate or advanced level of training that was provided to them during their undergraduate education, around 45% and 30% of them asked for advanced and moderate level of future training, respectively. It could be explained by the fact that they could have realized the significance of validity in their assessment practices. More specifically, when assessments are valid, they give information about learners' strengths and weaknesses in the language, and this could be used to tailor teaching approaches to meet learners' needs better. Also, ensuring that students' knowledge and skills in language are assessed exactly, results of assessment could be used to make reliable decisions regarding their language progress and proficiency.

Other two items which were “washback” and “authenticity” were among the areas for further training. Whereas nearly 60% of the participants reported to have received moderate to advanced level of training, an advanced level of future training need was expressed by nearly 40% of them. This result may be explained by the fact that these principles provide a framework for them in terms of creating and administering assessments which gauge learners' language level, progress, and achievements accurately. There could be some reasons why they need to receive more training in authenticity. Firstly, if assessment reflects real-world communication, students could be motivated to take more responsibility and enhance their ability to use language more autonomously, which are among the goals of language education. To put it in a different way, when they see the relevance of what they have learned to their daily lives, this makes learning more motivating and autonomous. Secondly, authentic assessment can enhance the validity of language tests in that if these assessments are authentic, they can give a better and more accurate picture of learners' language proficiency. This could be helpful for ensuring the reliability of the assessment results and making valid decisions about learners' abilities. These participants could have realized that receiving more training in “authenticity” could support

educational and assessment practices. Regarding “washback”, as teachers are conscious of the fact that the way language assessment is constructed and applied influences how students learn and how they teach significantly, they may need to deepen their understanding on this concept. Considering that the way of constructing and administering language assessment has an impact on what teachers focus on in their teaching practices, gaining more insight into this concept could amplify their teaching. More specifically, if assessments are intended to gauge only language knowledge like grammar and vocabulary, teachers could only concentrate on teaching these to students and may ignore language skills. If teachers are provided with more training in this pivotal concept, they could understand the effect of language testing on their practices better and they can take actions accordingly. Similarly, if the same exam that focused only on language knowledge but not skills is applied to students, they could focus only on these aspects rather than improving the overall proficiency. In that case, teachers’ awareness of the importance of washback plays a critical role in terms of designing assessments measuring an extensive range of skills in language, which will motivate and contribute to learners better in the long term. Therefore, their wish to improve themselves more in washback could possibly be explained by these.

As for “testing speaking in English”, which was the area in which the least amount of training was stated to have been received by the participants, most of them stated the further training need in this item. It could be explained by the fact that these participants could have realized that they lack adequate skills to assess speaking because of inadequate training offered to them. On the other hand, it was surprising to find out that though they wanted to develop themselves in testing speaking skill, “different test items/task types to test speaking in English” was among the least desired areas for extra training. Despite more than 10% of them revealing no prior training given to them in “different test items/task types to test speaking in English”, they did not want to improve equally as much. It could be attributed to some reasons. Firstly, they may not be aware of the latest

developments in testing speaking skill. Though they need to improve themselves in general terms of testing speaking, not wishing for learning various test types to test speaking could stem from their unawareness of these various types of assessment tools. It is of utmost importance to administer diverse forms of tests of speaking in accordance with various teaching contexts. Whereas sometimes interviews could be more suitable for testing learners' ability to communicate more formally, assessing students during group discussions could be better to gauge their skill in casual conversation. Therefore, without the knowledge and awareness of different ways of testing speaking, they could not measure the students' ability to speak properly.

The current study supports the results of previous studies (Ballidag & Inan Karagul, 2021; Hasselgreen, Carlsen, & Helness, 2004; Salami & Alharthi, 2022; Vogt & Tzagari, 2014) as need for extra training in this domain was uttered clearly. In the study of Ballidag & Inan Karagul (2021), participants perceived they need to receive training in this domain with various priorities for some aspects. The most desired areas for extra training were found for "testing and assessing aspects of culture" and "using statistics". Especially for testing culture, nearly one fourth of the participants stated that they did not take any training. Their desire for more training in this aspect could be related to this stated lack of training provided to them. As to the assessment of receptive and productive skills, there was also a wish for more training; however, the need for training in productive skills including writing and speaking was more than that for receptive skills involving reading and listening. When the level of training offered to them was considered for these two aspects, it was found out that the training level was the same. In spite of the same level of training received in productive and receptive skills, they wished to develop their knowledge more in the former aspect. This outcome is contrary to that of the current study in that while there is more need for training in receptive skills compared to productive skills in the present study, training need for productive skills were stated to be more than receptive skills in their study. It could stem from the difference in their teaching and assessment contexts, needs and goals of

learners besides strengths and weaknesses of teachers. Maybe, participants of the current study could consider themselves as more comfortable with assessing productive skills as compared with receptive skills. Therefore, it could be stated that priority for more training in certain areas could be related to different factors like sample group and their prior training levels. On the other hand, the least wished area for extra training in their study was “testing grammar” and “testing vocabulary”. This finding could be expressed by the fact that these respondents received the most training in these two aspects. They could have perceived the level of training they received as adequate, and they may not wish for more training in these micro-linguistic aspects of language. This is not consistent with the result of the current study in that in the current study there was a clear need for training with around 60% of them wishing for either advanced or moderate level of training need.

Salami and Alharthi’s (2022) study revealed that the participants needed to get more training in this domain, which is consistent with the current study’s outcome. While this domain was found to be the domain in which more training was stated to be received compared to the previous two domains, there was still need for extra training in their study. This shows that they do not regard their LAL level in content and concepts of LTA as adequate and wished to enhance this more. The area that they wished to get the most training was “testing integrated skills”. 46% of them perceived the level of training they received as basic and 40% of them needed advanced level of training as expected. As for the assessment of productive and receptive skills, they wanted to get more training in the former aspect, which is consistent with the outcome of the study of Ballidag and Inan Karagul (2021) and therefore contradicts with the outcome of the current study. Among the items of this domain, “testing receptive skills” was the least desired aspect for further training compared to the others. All in all, this present study and their study are similar in terms of need for further training in the last domain; however, there are differences in the priorities for the level of extra training need in different items of this domain.

Regarding the study of Vogt and Tsagari (2014), the domain of “content and concepts of LTA” was the area for which the most further training need was expressed compared to the other two areas. In the same vein, the present study demonstrated that the last domain was emphasized more for future training compared to the previous three domains. What is surprising is that despite the difference in sampling of both studies in terms of their profiles and prior training levels, both groups asked for further training. This demonstrates that this domain is perceived as considerable and must for their assessment practices and the improvement of LAL. In their study, some aspects of this domain are emphasized more for extra training than the others and when the finding for these items are compared to that of the current study, it was clearly seen that there were differences in the immediacy of more training need for some items. The top areas that were emphasized for more training need were “testing receptive skills”, “testing productive skills”, “testing micro-linguistic aspects”, “testing integrated language skills” and “testing culture”. More specifically, for testing these items, nearly 40% of them articulated advanced level of extra training despite the fact that almost the same percentage of them reported to have received advanced training. This could stem from the fact that they aim to develop their literacy more in these aspects and LAL level, too despite being offered adequate level of training. As these items including testing language skills separately such as speaking, listening, writing, reading, grammar, vocabulary and testing these language skills in an integrated way along with testing culture are critical for language learning and assessment practices, receiving training in these aspects will contribute to their LAL level overall and enable them to administer these practices better. With regard to “reliability”, “validity” and “using statistics”, they were the areas that were expressed to be ignored. Especially for “reliability” and “validity”, nearly 40% of them reported that they took no training and for “using statistics”, more than half of them (60.9%) mentioned no training was offered to them. Considering these neglected areas, they wished to improve their knowledge more as almost half of them asked for further advanced training. In this regard, the result accords with that of the current study as for these three areas there was a need for advanced training. Likewise, the study

of Hasselgreen et al. (2004) yielded similar findings in terms of prioritizing certain aspects including “assessing aspects of culture”, “assessing integrated skills”, “testing reliability”, “testing validity” and “using statistics” for future training needs. These areas were found to be conducted mostly without training and they asked for getting more education in these.

As stated by Vogt and Tsagari (2014), EFL teachers' assessment literacy in this domain was found as not developed. When the overall finding for received training for all items of this domain was analyzed, it was found that almost 30% and 35% of them stated no and a little training, respectively. Therefore, they found that nearly half of them stated advanced level for future training. They suggested that this could be due to their real-life experience and observations in the classrooms or requirements set by educational policies. It is mentioned that teachers are mostly engaged with testing and assessing language skills and knowledge rather than other aspects like testing culture or employing statistics. However, if they are not provided training in the vital aspects like testing culture or use of statistical analysis and these are ignored, it is stated to cause problems in overall competence and level of LAL. When teachers' wish for getting more training in these areas was analyzed, it could be concluded that these teachers do not want to disregard these aspects and they recognize the cruciality of them. All in all, the comparison of the present study and their study shows that this domain is highly critical for EFL language teachers, and they give priority to this domain overall apart from changing urgencies expressed for different items.

All in all, the second research question of the present study aimed to explore in-service EFL Teachers' needs for in-service training in language testing and assessment and the overall result for 4 domains demonstrated that the overall mean score was 1.72. Based on this, it is found that EFL language teachers perceive a need for future training at the moderate level. Considering their level of training that they received in undergraduate education was not totally adequate but nearly adequate, their desire for moderate level of training shows that they want to improve their LAL level and LTA knowledge overall. As put

forward by Vogt and Tsagari (2014), these respondents' level of LAL was developed to some extent but not totally. In spite of lacking very advanced level of LAL and LTA competency, they could be reported to have attained a certain level of LAL and LTA knowledge and they asked to develop this level more by receiving more training with priority being given to diverse aspects of testing and assessment. They could have considered that receiving intermediate level of training would be adequate for them to build upon their prior knowledge and level to attain a good level of LAL. There could be some factors for why they did not ask for advanced level of future training like time constraints, working conditions, resources or available training opportunities.

The overall finding accords with previous studies (Ballidag & Inan Karagul, 2021; Hasselgreen et al., 2004; Salami & Alharthi, 2022; Vogt & Tsagari, 2014) in terms of revealing the further need for receiving training in different domains of LTA. Particularly, in the current study the most desired areas for further training were "preparing achievement tests", "preparing placement tests", "preparing proficiency tests", "preparing progress tests", "preparing diagnostic tests", "using self/peer assessment", "using European Language Portfolio", "alternative assessment", "formative / summative assessment", "informal / formal assessment", "identifying what has been learned", "measuring general ability to learn a foreign language", "awarding final certificates", "testing integrated language skills", "practicality", "reliability" and "validity".

In the present study, the greatest amount of training need compared to other three areas was reported for "content and concepts of LTA", and the least amount of training need was found out for "classroom-focused LTA". When the level of training needs was considered in relation to the amount of training received by the same participants, it was somewhat surprising to find out that although "content and concepts of LTA" was the domain in which more training was reported to be received, they perceived the most training need for this domain. Similarly, though they reported to have received the lowest amount of training in "classroom-focused LTA", they also reported the lowest amount of training need

for the same domain. This is an unanticipated finding on the grounds that this domain would be expected to be the domain for which more training need would be expressed. In consistent with the current study, the study of Vogt and Tsagari (2014) also found out that the most desired area for future training was “content and concepts of LTA” despite its also being the area in which more training was stated to be provided to them. As highlighted by Vogt, Tsagari and Spanoudis (2020), there could be some factors at different micro-, meso- and macro-levels which influence their perceived level of future training needs. Their own classroom experience could be determinant for wish for extra training in LTA overall and especially more desire for training in certain domains like their instructional decisions, real assessment applications and conversations with students and their parents. Also, their institutions could play a role in their need for future training in LTA like communication and collaboration among teachers, training opportunities and the profile of school or language school in which they work. The other factor could be at the macrolevel including educational policies and culture of assessment. All of these factors could have also affected their desire for more training for certain domains and less training need for the other domains.

Discussion of the 3rd Research Question

Differences between EFL Teachers Regarding In-service Training Needs in LTA

The next research question of the present study was “Do in-service EFL teachers’ training needs in language testing and assessment vary according to (a) different educational settings, (b) the current level of education that they are teaching, (c) year of experience and (d) their education level?”. There were four groups within the respondents that formed the foundation for comparison: (1) the type of school or institution they work at, (2) the grade level they teach at, (3) level of experience, and (4) level of education. Data gathered from the questionnaire were analyzed to answer this question through one-way ANOVA and Tukey HSD (when needed).

In this study, there were 300 in-service EFL teachers who work at different formal and non-formal education institutions in Turkey. While 99 of them work at public schools, 108 of them work at private schools and 93 of them work at private language schools. The results of the test showed that in-service training needs showed no significant difference regarding these educational settings. Regardless of the type of institution they work at, their extra training needs in LTA were similar. There could be some underlying reasons why different educational settings do not make a difference in teachers' perceived training needs in LTA. It could be ascribed to the evolving nature of language testing and assessment field in that regardless of the type of institution they work, they are expected to develop sound assessment practices. Also, both formal and non-formal institutions expect them to have adequate level of knowledge and skills in assessment and apply it effectively, which could account for the similarity in the training needs of EFL teachers in both settings. The other reason could be related to macro-factors like education policies or opportunities for professional development. It is possible that such factors could have affected different types of educational settings in a similar manner, and this might have caused the extra training needs in LTA to be alike. It is critical to note that in-service training need of teachers is multifaceted, and it could be under the influence of diverse factors that are not related to different educational settings.

Another variable in the present study was the grade level at which they teach. Similar to the type of institution they work at, no significant differences were observed for in-service training needs of teachers working at primary, secondary and high schools. 102 of them work at primary school and 101 of them work at secondary school whereas 97 of them work at high school. The results of the test showed that teaching at different grade levels did not influence their need for further in-service training in LTA significantly. Despite the fact that they teach and assess their students at different grade levels, their wish for further training was found to be similar across teachers working at different grade levels. As the main goal of teaching language is to enhance and assess the acquisition of language

skills and knowledge and this is aimed at all levels of education, it is not surprising that the need for more training among EFL teachers did not differ from one another. At all levels of education, teachers need to have sound assessment practices and sufficient level of LAL despite differences in the learners' age group, proficiency levels and language learning and teaching goals. Thus, it is possible that the need for more training in this field does not alter with regard to these different grade levels. Moreover, it is crucial to point to the fact that maybe these teachers working at different levels of education might not have received adequate training in LTA. Their initial training could be limited in terms of LTA; therefore, they might seek to receive more training in this field to keep up with recent developments and knowledge regardless of different levels of education they work at. It is important to mention that despite this similarity in the need for future training in this field, the content of training may change based on the profile of the students like different age groups or proficiency levels they teach. While specific communicative assessment activities could be utilized with younger learners like in primary school, certain assessment methods like keeping portfolio or peer-assessment would be more suitable for high school students. All in all, these teachers' need for further training in LTA fields did not differ significantly across different levels of education.

When it comes to comparing the participants' needs in terms of their experience years, it was found out that in-service training needs showed no significant difference based on experience levels of teachers. In this study, 95 of them had 1-5 years of experience and 112 of them had 6-10 years of teaching experience while 93 of them had 11 and more years of teaching experience. Regardless of their year of experience in teaching, future training needs in LTA were similar. This finding was not surprising in that similar level of training was wished by EFL teachers no matter how many years of teaching experience they have had. It could be related to the multifaceted and ever-evolving nature of the language assessment field. The necessity for keeping up with recent developments in the field and updating their assessment knowledge as well as increasing their LAL level without

considering the level of teaching experience may be taken into account while interpreting similar level of training needs. In other words, no matter whether they are novice or experienced teachers, they feel the need for further training in these four domains of LTA. These teachers could also be stated to have the awareness that experience does not always ensure high level of LAL and competence in assessment practices. That's why, experienced teachers may have felt the need for further training similar to the novice teachers to enhance their LAL level. Moreover, it is possible that these experienced teachers were also aware that receiving more training in LTA contributes to being proficient assessors and testers, which can't be just guaranteed by having more teaching experience. The areas that they wished to get more training could change based on whether they are novice or experienced teachers in that while novice teachers could ask for more training in basic principles and concepts of assessment, experienced ones could desire for getting further training in different areas like utilizing large-scale assessments or carrying out validity studies, which could be also explored in future studies. All in all, to keep them updated and develop their current language assessment literacy levels, these EFL teachers uttered their need for extra training in LTA regardless of their experience levels.

On the other hand, as for their education level, there were 27 EFL teachers who hold MA and 12 holding PhD Degrees though there were 261 EFL teachers having BA degree. The results of the test revealed that in-service training needs showed significant difference regarding these education levels. When the results of Tukey HSD test were acquired, it was seen that the teachers with MA degree had more in-service training needs in LTA than teachers with BA degree. In other words, the in-service training needs in language testing and assessment field of EFL teachers holding MA and BA degrees were significantly different. No such finding was noted for other education levels. There could be some underlying reasons why holding MA and BA degrees make a difference in EFL teachers' perceived in-service training needs in LTA. Firstly, it is probable that these teachers who complete their MA degrees may ask for more training because they did not take LTA course

or the course they took was not adequate. As MA programs generally include second language acquisition, approaches, methods and techniques, they might not focus on English language testing deeply. When the curriculum of MA programs offered in Turkey was examined, just one language testing course was included in general. Secondly, EFL teachers having completed their MA degrees could be stated to have more awareness of the importance of level of LAL compared to those with BA degrees. In other words, those with an MA degree have completed a more advanced level of education and are likely to possess a more profound grasp of language teaching and learning practices including assessment. Maybe, they became more aware of the significance of assessment knowledge during the course offered to them in MA program. Also, they might have considered that they had expected to learn more from MA programs in terms of LTA field; however, they were offered limited course content and they may not be satisfied with their current level of knowledge or LAL level. They may have also found their BA education not adequate for providing this LAL level and expected MA to be more contributive, however, it might not have met this expectation. It could also result from their experience during their MA education or other factors like internal motivation or interest for inquiring more about LTA. Considering these, teachers with MA could have asked for more in-service training based on their experience and awareness raised in testing courses in MA programs. Furthermore, they might seek to receive more training to meet these necessities if certain teacher job vacancies oblige them to have increased mastery in LTA field, which will help them to attain professional growth and boost their career prospects.

As for teachers with PhD, they were found to have asked for in-service training in LTA less than the other two groups, but there was not a significant difference in their perceived training needs in LTA between teachers with PhD and the others. It could be ascribed to the possibility that EFL teachers having completed their PhD may have taken extensive courses or research in LTA field. Therefore, their desire for further training was less than the two other groups of teachers. However, this may not be the case all the time

in that some PhD programs do not offer additional testing courses, but rather they focus more on current issues in English Language Teaching, contemporary trends in Second Language Acquisition (SLA) and research like quantitative and qualitative research methods. All of them have to be considered in-depth while analyzing their perceived level of training needs in LTA.

All in all, needs for further in-service training in LTA field showed significant difference between EFL teachers with MA and BA degrees and aforementioned reasons may explain why teachers holding MA degree demand more training compared to those holding BA degree. Overall, other than the aforementioned variable which is education level, the statistical analysis revealed no significant difference for the participants' in-service training needs in language testing and assessment area.

Discussion of the 4th Research Question

In-service EFL Teachers' Experience Regarding LTA Practices in Classroom Context

The last research question of the current study was "How do in-service EFL teachers describe their experience in language testing and assessment practices in classroom context?". There were 11 open-ended questions which were grouped under 6 different categories. These groups were related to educational background of participants, ideas about Testing and Evaluation course offered to them in their undergraduate education, application of LTA, training needs in LTA, problematic issues in LTA and solutions to these problems and further comments. The first three questions were about their educational background including their departments from BA, MA and PhD, year of experience in teaching and the number of LTA and statistics courses that they took. When these were analyzed, it was found that 14 out of 20 language teachers have BA degree in ELT and ELIT (10 and 4, respectively) and 6 of them hold MA degree in ELT and Curriculum and Instruction. Furthermore, their experience level differs from 1-5 years to 11 or more years and 9 of them have more than 11 years of experience whereas 5 of them have 6-10 years

and 6 of them 1-5 years of teaching experience. When the institutions and teaching levels they teach at are analyzed, it was found that they work at diverse levels in different schools including state, private and private language schools. With regard to the number of courses, all of them mentioned 1 LTA course. 3 of them also stated that they were provided with additional statistics courses.

The fourth interview question was related to whether the course in LTA that they were offered in undergraduate training covered essential constituents of LTA or not and majority of them stated that they did not evaluate this course enough for them. More specifically, 15 out of 20 participants stated it to be inadequate while 5 of them specified it to be enough. Among the ones who claimed that the course covered important aspects of LTA, few of them articulated that they received adequate amount of training in specific topics like how to prepare and score exams and providing feedback to learners based on the results of these exams. Also, these language teachers expressed that they do not feel the need for further training on the grounds that they could also learn and explore issues related to testing in their profession as they gain experience. This could be attributed to their belief in that as they believed they learned the essentials of testing, they could build upon their knowledge through practice. Also, it could be related to the fact that they regard their practice in LTA as adequate and they do not feel difficulty in utilizing tests. Especially, one of the participants enunciated he could employ language tests and exams without putting much effort and being expert. This could have led him to regard the training to be adequate. The other participant attributed the course adequacy to their professors in undergraduate training and specifically stated that since their professors were competent in LTA and taught them theoretical and practical aspects of the field. As these findings reveal, it could be stated that the ones who regard the course to have covered important aspects of LTA attribute this to some factors like belief of exploring other aspects in the profession itself, not having challenge in practice and faith in the expertise of their instructors in the university education.

However, the majority of them pronounced that Testing and Evaluation Course did not successfully include all the required components of LTA by providing some reasons. Most of them expressed that they noticed that inadequacy when they started teaching. Once they started teaching and testing practice, they came to realize that they should have taken more training in the assessment field. Particularly, one of them clarified the issue by mentioning that she noticed the necessity of being offered more training for the aim of being creative and motivating in test practices. It could have resulted from the fact that when they applied testing, they generally had the tendency to use the same or similar type of testing rather than innovative and alternative ways, which could make them feel that they have limited knowledge. As for the other participant, he mentioned that he does not seek to be expert, but be better in assessment practices, which he realized after graduating from the university. The other language teacher highlighted another dimension in terms of referring to the comprehensiveness of assessment field. She expressed that as this field not only covers measuring students' knowledge but also necessitates giving feedback and making contribution to their learning and development, she did not evaluate their training to be efficient enough to include all these aspects. Moreover, she added that she wished for more courses in this field in undergraduate education.

All in all, the analysis of the 20 in-service English language teacher interviews showed that Testing and Evaluation Course that was offered in undergraduate training did not incorporate required aspects of the assessment field. When the results of questionnaire and interview results were considered together regarding the adequacy of the course offered in undergraduate education, it was found that the results obtained from questionnaire and interview analysis showed similarities.

In the fifth interview question, in-service EFL language teachers were asked to list the most beneficial topics that were included in the course for them, and 8 items were uttered by them, which were alternative assessment, validity, how to adapt prepared tests, practicality, informal/formal assessment, objective testing, formative/summative

assessment, and testing reading. They mentioned that these were so useful for them in their classroom practice.

One of the interviewees remarked alternative assessment by focusing on the necessity for staying updated with the new generation learners in that instead of traditional ways of assessment, applying new alternatives in assessment like giving place to portfolios and presentations as well as self-assessment are better for all assessors and assessees. These forms of assessment would provide more learner and process- focused assessment rather than product- and result focused approach. There could be several possible explanations for this. Firstly, the new generation of students are accustomed to interactive learning and technology-based platforms. Using alternative assessments like project-based assignments or individual or group presentations could be more motivating for them in comparison to traditional tests like pen-and-paper exams. This would enhance their interest in the topic and promote deeper and permanent learning. Secondly, such types of assessment provide learners with applying their knowledge and abilities in real-world contexts like problem-solving tasks or simulations, and critical thinking and problem-solving skills of them will be encouraged. Thirdly, teachers can be stated to be aware of multiple intelligences and diverse learning styles in that while traditional assessments generally measure a narrow range of skills and mainly prioritize linguistic and logical intelligences, alternatives cater to a wider spectrum of intelligences and learning styles like visual or interpersonal intelligences. Moreover, using alternative assessments promotes collaboration and communication among learners. Group projects, debates, or interactive discussions can develop teamwork, interpersonal skills, and effective communication abilities, which are vital in today's collaborative work environments. Another reason could be the fact that they can tailor this type of assessment to individual students' needs, allowing for personalized and differentiated assessment strategies. This approach acknowledges the diverse learning profiles and accommodates different levels of proficiency, promoting inclusive education. Lastly, they offer a more authentic representation of students' abilities

on the grounds that they mirror real-life situations and tasks. This can provide a more comprehensive view of their skills, knowledge, and readiness for future challenges or careers, which is beyond the limited scope of traditional tests. All in all, these interviewees could be stated to be conscious of the fact that diverse types of assessments could complement each other and provide a well-rounded assessment approach for the new generation of students.

The other interviewees also stated informal and formal assessment in reference to the new generation of students, as well. Several factors could explain this. First of all, they could have considered that formal assessments, such as standardized tests or exams, often provide a standardized and structured way to measure students' knowledge and skills. However, they may not capture the full range of a student's abilities or provide insights into their individual progress. Informal assessments, on the other hand, can offer a more holistic view of students' learning by considering their everyday performance, observations, discussions, and interactions. Combining both formal and informal assessments allow for a more comprehensive understanding of students' strengths, weaknesses, and progress. Secondly, incorporating both formal and informal assessments will enable them to adopt a more individualized approach. Formal assessments provide a standardized baseline, while informal assessments enable educators to tailor instruction and provide targeted support to meet the unique needs of each student. Thirdly, traditional tests often take a longer time for grading and providing feedback, which could hinder students' ability to reflect on their performance and make improvements. Informal assessments, such as classroom discussions, formative assessments, or feedback during project work, offer more immediate feedback and allow for timely interventions and support. This helps students to address misconceptions, revise their work, and make progress in a timely manner. Besides, formal tests may sometimes sound disconnected from real-world applications and contexts. Informal assessments; on the other hand, such as projects, presentations, or portfolios, provide opportunities for students to engage in authentic learning experiences that mimic

real-world situations, which might enhance their critical thinking, problem-solving, and communication skills, preparing them for the challenges they will face outside of the classroom. Last but not least, they could have considered this as useful for holistic development of learners in that they require a broader focus on holistic development, including social-emotional skills, creativity, and critical thinking. Informal assessments allow for the evaluation of these skills and dispositions, which are not easily captured through traditional tests. This comprehensive assessment approach ensures that students' growth in all areas is acknowledged and supported. Incorporating both formal and informal assessments will help assessors to gather a more complete and nuanced understanding of students' abilities, promote personalized learning, provide timely feedback and support, and foster the holistic development of the new generation of students.

For the other items like validity and practicality, they highlighted these principles' importance. To start with validity, this could be attributed to some factors like informed decision-making, fairness and equity and accurate assessment. Valid assessments enable language teachers to make informed decisions about instructional strategies, curriculum development, and interventions. By understanding the validity of assessment tools, teachers can trust the results and use them as a basis for planning targeted instruction and supporting students' language development effectively. Also, valid assessments ensure fairness and equity by measuring language abilities impartially, without bias or discrimination. Language teachers need to ensure that their assessment practices are valid, allowing all students, regardless of their background or characteristics, to demonstrate their true language skills. As they also ensure that the results reflect students' true language proficiency and abilities. Language teachers need to use valid assessment methods to obtain reliable information about students' language skills, knowledge, and progress. These teachers could consider validity as useful for identifying the strengths and areas which need improvement precisely. As to practicality, their statement for its usefulness could be related to several aspects including efficiency and time management, resource allocation, teacher

autonomy, student engagement and alignment with learning objectives. Practical assessments are efficient in terms of administration, scoring, and feedback and EFL language teachers often have limited instructional time, so practical assessments allow them to gather meaningful information about students' language abilities without placing excessive burden on instructional time or administrative tasks. Also, they could have considered the availability of resources, such as materials, technology, and human resources. They can design assessments that align with the available resources, ensuring effective implementation and avoiding unnecessary strain on limited resources. Another benefit could be autonomy that practical assessments offer to teachers in that they provide language teachers with the flexibility to adapt and modify assessment methods according to their instructional context and learners' needs. Teachers can choose assessment strategies that are manageable and suitable for their specific teaching environment, ensuring a seamless integration of assessment into their overall instructional practices. The other benefit could be engaging students more since assessments that are practical and manageable for students to complete help to keep their interest, enthusiasm, and investment in the language learning process. Lastly, practical assessments align with the desired learning outcomes and instructional goals. Teachers can design assessments that directly assess the targeted language skills and knowledge, ensuring that the assessment results provide relevant information for instructional decision-making. Considering these factors, these interviewees could have stated the utilities of these two principles in their classroom experience.

As for objective testing, these interviewees could have regarded this component as fruitful because of diverse reasons. One possible explanation could be standardized evaluation as objective testing provides a standardized and consistent method of evaluating students' language proficiency. It ensures that all students are assessed based on the same criteria and that the assessment is fair and impartial. This objectivity ensures fairness in the assessment process, treating all students equally and reducing potential biases that may

arise from subjective assessment methods. Additionally, this type of testing, especially through multiple-choice questions or other structured formats, tends to yield reliable results. The scoring process is typically clear-cut and less susceptible to subjective interpretation. This reliability enables teachers to compare students' performance, track progress over time, and make informed decisions based on the assessment outcomes. This could also help them as it is often efficient and time-saving in terms of administration and scoring similar to practicality. Furthermore, they can assess a large number of students relatively quickly, which is especially beneficial when dealing with large class sizes or limited time for assessment. As objective testing methods are commonly used in external language proficiency exams, familiarizing students with objective testing formats in class assessments can help them become more comfortable and confident when taking external exams. It familiarizes students with the types of questions they are likely to encounter, the time constraints, and the expectations of objective assessment. Also, as objective tests often provide clear feedback to students, indicating which items they answered correctly or incorrectly, this feedback can help students identify areas of strength and areas that require improvement. Language teachers can use this feedback to guide instructional interventions and address specific language learning needs. The other benefit could be opportunity of comprehensive coverage of objective testing in that a variety of language skills and content can be integrated, and this will give a holistic view of students' language proficiency and enables teachers to assess various language components effectively. Other advantages of this type of testing can be alignment with standards in that it ensures assessment outcomes are relevant and useful for evaluating students' progress in relation to established benchmarks.

Moreover, knowing about formative and summative testing is highly effective for language teachers in their assessment practices for several reasons. It is highly possible that through incorporating both formative and summative testing in their assessment practices, language teachers can gather comprehensive information about students'

progress, provide timely feedback, differentiate instruction, monitor teaching effectiveness, ensure accountability, and support students' ongoing language development. The combination of these two types of assessment contributes to a balanced and holistic approach to evaluating student learning outcomes. More specifically, formative testing allows language teachers to gather data on individual students' progress, enabling them to differentiate instruction based on students' specific needs. Teachers can tailor their instructional approaches, provide targeted interventions, and offer additional support to students who require it, ensuring that all students have the opportunity to achieve success. It also provides students with timely and specific feedback on their performance, highlighting areas for improvement. Language teachers can offer constructive feedback, suggest strategies for enhancement, and engage students in reflective practices. This feedback loop promotes self-assessment and self-regulation skills, encouraging students to take an active role in their language learning journey. On the other hand, summative testing occurs at the end of a learning period and evaluates students' overall language proficiency or mastery of specific content or skills. It provides a comprehensive snapshot of what students have learned and achieved. Summative testing allows language teachers to determine the extent to which students have met the learning objectives or standards set for the instructional period. It also plays a crucial role in evaluating students' performance and assigning grades or marks. It provides a formal and systematic way to assess students' language abilities, ensuring accountability and comparability across students. Summative testing results are often used for reporting purposes, such as progress reports or end-of-term evaluations. Considering these utilities of both functions of assessment, in-service language teachers could have mentioned this to be effective for their classroom practices.

The other component stated to be helpful for them was testing reading. It could stem from the fact that these language teachers are aware of the importance of reading skill and assessment of it. Especially, one of them specified that they were taught diverse test items to assess this skill such as interactive and extensive reading. Also, she

wished that training was continued to be given in these in teacher education programs. The highlight for testing reading could be linked to some factors. As reading tests allow teachers to gauge students' understanding of written texts. By evaluating their ability to comprehend and interpret the content, teachers can identify areas where students may struggle and provide targeted instruction or support. By analyzing their current reading abilities, they can establish targets for improvement and design instructional activities that align with these goals. This process helps students stay motivated and provides a clear path for their language development. Reading assessments also offer an opportunity for teachers to provide timely and constructive feedback to learners. It helps them to identify areas of improvement and provide specific feedback on their reading strategies, comprehension skills, vocabulary usage, and more. This feedback guides students' learning and helps them make progress. As for the variety of test items, they enable teachers to gather a comprehensive picture of students' reading abilities. Each test item provides a unique perspective on students' skills and knowledge, contributing to a more accurate assessment. For instance, tasks such as matching headings to paragraphs or completing a graphic organizer based on the text require students to interact with the material actively. This active engagement promotes deeper comprehension and helps teachers assess students' ability to organize and synthesize information. The other utility is related to accommodating diverse styles of learning in that including different test items caters to these differences and provides a fair assessment for all students. For instance, some learners may excel in multiple-choice questions, while others may prefer tasks that require short responses or extended written answers. By using varied test items, teachers can provide opportunities for students to demonstrate their strengths and abilities in different formats. Their emphasis for testing reading and different test items to test this skill could stem from these reasons.

The last item that was itemized by interviewees was the way to adapt prepared tests. One of the interviewees laid emphasis on it by mentioning that they were not experts of testing; hence, they need to adapt ready-made tests considering the reliability issue of

testing. He emphasized the need for information of criterion for adapting ready materials like adapting the length or simplifying the test items. The other participant also mentioned the difficulty of preparing a test from scratch and the knowledge of how to adapt existing tests to their contexts was very worthwhile. There could be other underlying reasons for their preference like a practical and efficient approach to assessment, saving time, maintaining consistency, and focusing more on instructional planning and support for their students rather than spending much time in preparing tests. As they suggested, adapting ready-made tests assists to maintain the validity and reliability of the assessment. Modifying or adapting existing test items ensures that they align with the specific objectives and context of the classroom, while still retaining the core validity and reliability of the original test. This helps ensure that the assessment accurately measures what it intends to measure. Besides, it frees up time and cognitive load for teachers, allowing them to focus more on instructional planning and tailoring instruction to students' needs. By utilizing existing tests, teachers can spend more time analyzing assessment results, identifying areas for improvement, and designing targeted interventions or differentiated instruction.

Overall, the analysis of the fifth interview question put forward that alternative assessment, validity, how to adapt prepared tests, practicality, informal/formal assessment, objective testing, formative/summative assessment and testing reading were found to be very useful for them in real classroom practices for a great many reasons.

The sixth interview question was related to which topics they would like to learn more if they took this course again or took an advanced course in language testing and assessment. They listed certain subjects which encapsulated statistics, alternative assessment, computer-based testing, reliability, approaches to language testing, giving feedback, testing integrated language skills and validity.

To start with alternative assessment and validity, which were also stated to be useful topics as discussed in the previous part, these participants emphasized these topics in case of the fact that they took the same course in LTA again or they would take any other course

in the future. These two topics were consistently pronounced, and this could be explained by the fact that these in-service language teachers gave importance to these components. Alternative assessment methods are valuable for EFL teachers as they promote a more comprehensive evaluation of teachers' abilities to assess learners effectively. By utilizing alternative assessment, teachers can demonstrate their competence in designing and implementing varied assessment techniques that align with principles of authenticity, validity, and fairness. It encourages teachers to think creatively and apply innovative assessment approaches that are more suitable for assessing language learning and development. Besides, validity allows them to have confidence in the assessment process and results. It ensures that the assessment accurately reflects their knowledge and skills in assessment practices, which in turn supports their professional development and growth as EFL educators.

This could also be discussed in light of the results of questionnaire in that “alternative assessment”, “validity”, “reliability” and “testing integrated language skills” were among the items with high mean scores revealing the need for further training. Especially, for these items, majority of the participants stated either intermediate or advanced level of training. In this regard, this finding accords with the previous finding of the current study, which was of utmost importance considering the future training needs of language teachers and designing further courses in this field.

Language teachers may benefit from more training in reliability and testing integrated language skills for these reasons. As reliability ensures that an assessment instrument or process produces consistent and dependable outcomes when used repeatedly, training in reliability helps language teachers understand how to design and administer assessments that yield consistent results. By enhancing their knowledge of reliability, teachers can ensure that their assessments are trustworthy and free from measurement errors or inconsistencies. Reliable assessments provide more accurate and consistent feedback on students' language skills, enabling teachers to make informed

instructional decisions. It also promotes fairness and equity in the evaluation process, as students can have confidence that their performance is being consistently and reliably assessed.

As for testing integrated language skills, the ability to effectively use multiple language components (reading, writing, speaking, and listening) together in a cohesive and communicative manner is significant for language teachers. Thanks to this, they could evaluate students' proficiency in applying language skills across various contexts. Moreover, language teachers need training in testing integrated language skills to design assessments that align with real-world language use. Integrated assessments capture students' ability to apply their language skills holistically, reflecting how language is used in authentic communication. It helps teachers evaluate students' language proficiency more accurately, as opposed to assessing skills in isolation. Additionally, training in this area equips teachers with strategies for developing integrated tasks, grading rubrics, and effective feedback mechanisms that foster students' overall language development. By receiving training in reliability and testing integrated language skills, teachers can improve their assessment practices, ensure the consistency and accuracy of evaluations, and provide more comprehensive feedback to students. These skills contribute to creating a more valid, reliable, and fair assessment environment, supporting students' language learning and overall proficiency.

As for their wish for training in computer-based testing and approaches to language testing, these allow language teachers to stay updated with technological advancements and assessment methodologies. It equips them with the necessary skills to design and implement computer-based assessments effectively and to align their assessments with sound language testing approaches. Ultimately, this training enhances teachers' ability to create valid, reliable, and pedagogically sound assessments that support students' language learning and proficiency. With the increasing integration of technology in education, computer-based testing has become more prevalent in language assessment.

Training in computer-based testing equips language teachers with the skills and knowledge necessary to effectively administer, score, and interpret assessments conducted using computer-based platforms. They could be stated to be aware of various advantages of it such as automated scoring, immediate feedback, enhanced test security, and flexibility in test administration. By receiving training in this area, teachers can harness the benefits of technology to create engaging and efficient assessments. They can learn how to design interactive tasks, utilize multimedia resources, and effectively analyze computer-generated test data. This training empowers teachers to leverage technology for more efficient and reliable language assessment practices. As for training in approaches to language testing, it familiarizes teachers with various assessment models, such as communicative language testing, formative assessment, performance-based assessment, or criterion-referenced assessment. Understanding different approaches to language testing enables teachers to align their assessments with their instructional goals and promote effective language learning. It helps teachers design assessments that assess not only linguistic knowledge but also students' ability to use language in authentic contexts. They can create assessments that capture the complexity of language skills, promote critical thinking and problem-solving, and provide meaningful feedback to students by using appropriate assessment approaches. This training empowers teachers to develop assessments that support students' language development and reflect current best practices in language assessment.

The finding of need for more training in these two components was unexpected since “approaches to language testing” had the lowest mean score in its domain and it was followed by “computer-based testing” according to the questionnaire findings. Nearly 10% of them stated no need for extra training in the questionnaire; however, in the interviews, they were emphasized for future topics to be covered. This discrepancy could be related to some factors like awareness and perception, perceived relevance, professional development and growth. It is possible that language teachers may not have been fully

aware of the specific concepts and techniques related to approaches to language testing and computer-based testing. As a result, their mean scores in these domains might have been lower. However, during the interviews, when prompted with specific questions or topics related to these areas, teachers may have realized the importance and relevance of gaining further knowledge and skills in these domains. Their emphasis on these topics during the interviews may reflect a growing awareness and perception of the significance of these training areas. The discrepancy could also be attributed to the perceived relevance and immediate applicability of the topics. While the mean scores in the questionnaire may have been lower, teachers might have recognized the practical value and benefits of acquiring knowledge and skills in approaches to language testing and computer-based testing. They may have emphasized these topics in the interviews based on their understanding of how these skills can positively impact their teaching practice and enhance student learning outcomes. Moreover, it is usual for them to realize their own learning needs and areas for improvement through reflective processes, such as interviews or discussions. The interview process could have provided an opportunity for teachers to reflect on their own practices and identify areas where they could benefit from further training. The emphasis on these topics in the interviews might reflect their desire for professional development and growth in these specific areas. Shortly, the unexpected finding can be attributed to factors such as awareness, the perceived relevance of these topics and professional development needs. The interview process likely provided a deeper exploration of teachers' perspectives, leading to a more nuanced understanding of their training needs and priorities.

Another item that was mentioned was “giving feedback” and language teachers could have stated this component for future training for various reasons. First of all, feedback plays a crucial role in promoting student learning and growth. Effective feedback provides students with specific information about their language performance, highlights areas of strength and areas needing improvement, and offers guidance on how to enhance

their language skills. Training in giving feedback equips teachers with the knowledge and strategies to provide constructive and actionable feedback that supports students' language development. Secondly, enables teachers to tailor their feedback to individual students' needs. It helps teachers develop the ability to provide personalized guidance that addresses specific areas of improvement for each student, ultimately promoting more targeted and effective instruction. Moreover, it helps teachers develop the ability to communicate feedback clearly, respectfully, and in a way that fosters a positive and supportive learning environment. Teachers learn techniques for providing feedback that is specific, timely, and understandable to students, promoting effective communication between teachers and learners. The other reason could be the fact that it cultivates reflective practice as it encourages them to engage in reflective practice. It prompts teachers to reflect on their own teaching strategies, student performance, and the impact of their feedback. By analyzing the effectiveness of their feedback, teachers can continuously improve their instructional practices and adapt their feedback techniques to better meet the needs of their students. Ultimately, their wish for extra training in giving feedback could be related to equipping themselves with the skills and strategies to provide constructive and meaningful feedback that enhances student learning, promotes individualized instruction, motivates students, fosters effective communication, and cultivates reflective teaching practices. It empowers teachers to be more effective in their roles as facilitators of language learning and supports students in their language development journey.

Regarding the use of statistics, this is also among the topics that were wished to be covered again. This could have stemmed from a great number of reasons. They could wish for more training as training in statistics helps language teachers design, develop, and evaluate language assessments more effectively. Statistical knowledge enables teachers to analyze item difficulty, item discrimination, reliability, and validity of assessment instruments. It allows them to make informed decisions about the quality and appropriateness of test items, leading to more reliable and valid assessments. Also,

statistical skills empower language teachers to analyze assessment data accurately and interpret the results meaningfully. They can use statistical techniques to identify trends, patterns, and areas for improvement in student performance. By understanding data analysis, teachers can identify specific language learning needs, adjust their instructional strategies, and provide targeted interventions to support student progress. In addition to these, statistical literacy enables language teachers to engage with research studies and educational data. By understanding statistical concepts and methods, teachers can critically evaluate research findings and apply evidence-based approaches to their teaching practice. They can use statistical evidence to inform their instructional decisions, assess the effectiveness of different teaching strategies, and make data-driven adjustments to improve learning outcomes. The other reason could be that they can employ statistical techniques to evaluate the reliability of assessments and make improvements to enhance their validity. Moreover, it could allow teachers to actively engage in professional development activities and collaborate with other professionals in the field. Statistical literacy helps teachers communicate effectively with assessment specialists, researchers, and colleagues, fostering a collaborative environment for continuous improvement in language assessment practices. In short, training in statistics in language assessment equips language teachers with the skills to develop reliable and valid assessments, analyze assessment data, make evidence-based decisions, and engage in professional collaborations. It enhances their ability to create effective assessments, interpret data accurately, and improve their teaching practices based on empirical evidence. Ultimately, this training contributes to the enhancement of language assessment practices and supports teachers in fostering better language learning outcomes for their students.

All in all, the analysis of the sixth interview question suggests that some topics which were statistics, alternative assessment, computer-based testing, reliability, approaches to language testing, giving feedback, testing integrated language skills and validity were

wished to be included in LTA course if they were to take any advanced course or took the same course again for various reasons.

In the seventh interview question, in-service EFL language teachers were asked to mention the way they decided which assessment method to employ in their classrooms. When their responses were analyzed, it was found out that there were some factors that determine their choice of assessment methods to be utilized like talking to colleagues, using the internet, checking books in testing and assessment and using previous knowledge or practice.

To start with the first common factor, which is talking to colleagues, they considered this way as practical and easier. Also, they highlighted the phenomenon of cooperation and feeling more relaxed while consulting their colleagues. There could be several reasons for this and certain advantages of seeking input from their workmates. First of all, colleagues who have been teaching for a longer time or have experience with different assessment methods can provide valuable insights. They may have tried various approaches and can share their successes, challenges, and lessons learned. This sharing of knowledge helps teachers avoid reinventing the wheel and benefit from the collective wisdom of their colleagues. As one of the participants stated, their workmates are familiar with classroom or school context along with existing realities and challenges, which helps them in their decision. Secondly, colleagues may have different teaching styles, backgrounds, or expertise, leading to a variety of viewpoints on assessment. Consulting them provides teachers with a broader range of ideas and strategies. Different perspectives can spark creativity, inspire innovation, and encourage teachers to consider alternative approaches they may not have initially thought of. Moreover, engaging in discussions with colleagues about assessment methods fosters professional growth. It allows teachers to expand their knowledge, learn about new trends and research in language assessment, and stay updated on best practices. By collaborating with others, teachers can improve their skills and enhance their teaching effectiveness. Another advantage can be that seeking input

from colleagues can help ensure the quality and validity of assessment methods. Colleagues can review and provide feedback on proposed assessments, helping to identify potential biases, flaws, or areas for improvement. This collaborative process promotes the development of fair, reliable, and valid assessments that accurately measure students' language proficiency. The other benefit could be that by consulting colleagues, language teachers create a supportive network where they can share their concerns, seek advice, and gain reassurance. Collaborative decision-making fosters a sense of rapport and solidarity among teachers, leading to a more positive and enriching professional environment. Ultimately, the main goal of assessment is to support student learning. By consulting colleagues, language teachers can select appropriate assessment methods that align with their instructional goals and cater to the needs and abilities of their students. Engaging in reflective discussions and receiving feedback from colleagues ensures that assessments are effective in measuring student progress and providing meaningful feedback for improvement. In summary, consulting colleagues about assessment methods in language classrooms brings numerous advantages, including the sharing of expertise, diverse perspectives, professional growth, quality assurance, support, and improved student learning outcomes. Collaboration among teachers enhances the overall teaching and assessment practices, benefiting both educators and students alike.

As for the second factor, which is using the internet to determine which assessment method to use in classroom, interviewees stated that the internet allows them to find different assessment ways and opinions to use. Several advantages could explain their preference of using the internet. As they stated, they could access a wide range of resources related to assessment. The internet provides language teachers with a vast array of resources related to assessment methods. They can find articles, research papers, blogs, teaching websites, and educational platforms that discuss various assessment strategies and techniques. This access to a wealth of information allows teachers to explore different options and stay updated on the latest trends and developments in language assessment.

As one of them mentioned, discovering new ways of assessment prevents monotonousness among learners and themselves, too. Also, through the internet, language teachers can connect with educators from around the world. They can participate in online forums, join professional communities, and engage in discussions with teachers who have different cultural backgrounds and teaching contexts. This global input provides a rich diversity of perspectives on assessment methods, offering teachers a broader range of ideas and insights. The other benefit could be that the internet allows language teachers to access information about assessment methods at their convenience. They can search for resources, read articles, and watch videos at any time and from anywhere. This flexibility enables teachers to engage in professional development and research without being limited by geographic or time constraints. Besides, online platforms often provide case studies, examples, and practical applications of assessment methods. Language teachers can find real-world examples of how particular assessment techniques have been implemented successfully in various educational settings. These resources offer insights into the practical aspects of using different assessment methods, helping teachers make informed decisions based on the experiences of others. In line with the first factor, it also facilitates collaboration and networking among language teachers. Online platforms, forums, and social media groups dedicated to language teaching and assessment allow teachers to connect, share ideas, and learn from each other. They can discuss assessment methods, exchange resources, seek advice, and provide support to fellow educators. This collaborative environment fosters professional growth and promotes a sense of community among language teachers. Last but not least, due to the fact that language assessment is a dynamic field with new research and approaches emerging regularly, using the internet is advantageous for them as it provides a platform for teachers to access the most up-to-date information on assessment methods. They can access research findings, attend webinars, and follow experts and organizations that specialize in language assessment. This helps teachers stay informed about the latest advancements and evidence-based practices. In short, they could have used the internet for many reasons including easy access to a wide

range of resources, diverse perspectives, updated information, practical examples, and opportunities for collaboration. By leveraging online platforms, teachers can enhance their knowledge, discover innovative assessment methods, and connect with a global community of educators. The internet has become an indispensable tool for researching, exploring, and staying informed about assessment methods in language classrooms.

The other factor that was preferred by them was looking at books in the LTA field and they mentioned that this way made them feel more assured. There could be certain advantages of referring to books and literature on assessment methods to inform their decision-making process. Firstly, as emphasized by interviewees, these resources offer a comprehensive overview of various assessment methods, including their principles, techniques, and practical applications. Teachers can benefit from the expertise of assessment experts and researchers who have studied and analyzed different approaches. This could also be seen in the comments of some interviewees in that looking at theoretical books in the field ensures them with their practices and allows them to feel more secure thanks to the expertise and valid knowledge in the books. In addition, books on assessment methods typically incorporate research findings and evidence-based practices. They provide insights into the effectiveness and validity of different assessment techniques, helping teachers make informed decisions about which methods are most suitable for their specific teaching context. By relying on evidence-based practices, teachers can ensure that their assessments are reliable, valid, and fair. Also, some assessment books often offer practical guidance and examples of how to implement specific assessment methods. They provide step-by-step instructions, sample tasks, and scoring rubrics that teachers can adapt and use in their classrooms. These resources help teachers understand the practical aspects of assessment and offer ideas for designing and implementing assessments effectively. The other benefit might be related to the alignment with curriculum, objectives and standards. They can receive guidance on how to design assessments that measure students' progress in relation to specific learning outcomes. By referring to these books,

teachers can ensure that their assessments align with their teaching goals and accurately measure student achievement. Teachers also can deepen their understanding of assessment principles, explore new ideas, and refine their assessment practices through independent reading and study. They serve as reference materials that teachers can consult whenever they encounter challenges or need inspiration for designing assessments. They also offer a wealth of knowledge and support for teachers at various stages of their teaching careers. In short, they could have referred to this factor as a way to make a decision in the methods to apply during assessment practices considering a range of benefits including established frameworks and theories, evidence-based practices, practical guidance, alignment with curriculum and standards, opportunities for professional development, and comprehensive resources. By referring to these books, teachers can enhance their knowledge, refine their assessment practices, and ensure the effectiveness and fairness of their assessments in language classrooms.

The last factor addressed in the comments of participants was drawing upon their prior knowledge and practices. A group of interviewees evaluated this method as more effective and useful. They believed that experience is a valuable tool for facilitating learning and reflecting upon this experience is the best way to determine assessment method. Some possible explanations could account for this. In the first place, teachers' prior knowledge and experiences with certain assessment methods provide them with a level of familiarity and confidence. They have firsthand experience implementing these methods and are aware of their strengths, limitations, and potential adaptations. This familiarity allows teachers to feel more comfortable and competent when selecting and implementing assessment techniques. In the second place, their prior knowledge and practices in assessment are closely tied to their teaching style and instructional objectives. They understand how different assessment methods align with their specific teaching approaches and goals. By utilizing their prior knowledge, teachers can choose assessment methods that complement their teaching style and effectively measure the desired learning

outcomes. To exemplify, one in-service language teacher particularly stated that the assessment method should be in harmony with her teaching philosophy and style by giving an example related to process or product-oriented assessment. If there was a mismatch between the style and application, this would create problems for both learners and instructors. Furthermore, their prior knowledge and practices provide them with a repertoire of strategies and techniques that can be tailored to address the unique characteristics and requirements of their learners. This adaptability allows teachers to customize assessments and make them more relevant and meaningful for their students. Another possible factor is that reflecting on past experiences and outcomes helps teachers refine their assessment practices. By drawing on their prior knowledge, teachers can assess the effectiveness of previous assessment methods, identify areas for improvement, and make informed decisions about which methods to use in the future. This reflective process contributes to professional growth and the continuous improvement of assessment practices. Besides, they have a repertoire of tried-and-tested methods that have worked well in the past, allowing them to save time and effort in designing assessments. This efficiency allows teachers to focus more on the quality of assessments and the analysis of student performance. For instance, one of them signified that considering previous assessment applications that they used helps them in that whether they are applicable and effective for testers and testees, as well.

The other benefit could be that by relying on their prior knowledge and practices, teachers can maintain consistency and continuity in assessment across different learning contexts. They can apply consistent standards, scoring criteria, and feedback approaches that have proven effective. This consistency ensures fairness and comparability in assessing student progress over time. All in all, language teachers' prior knowledge and practices in assessment bring advantages such as familiarity, alignment with teaching style, adaptability, reflection and improvement, efficient planning, and consistency. By leveraging their own experiences and expertise, teachers can make informed decisions and design

assessments that best meet the needs of their students and align with their instructional objectives.

The eighth question in the interview was whether LTA training in their university education is sufficient for them and if they think the training is not adequate, they were asked to list the topics that they require more training. Only 6 out of 20 participants evaluated this as sufficient and the rest of them did not find the training adequate. This was in line with the findings of the questionnaire that revealed participants regarded education not completely adequate, but partially. Analysis of both questionnaire and interview reveals clearly that LTA training offered in undergraduate education is not adequate according to these in-service EFL teachers. As for the topics that were mentioned commonly, they were statistics, ELP and testing grammar.

In the first place, “statistics” was mentioned as stated before in the previous interview question that was related to the topics that they wished to be covered again in any assessment course. Considering this frequency, it could be clearly pointed out that these language teachers need to take training in the use of statistics strongly. In other words, they prioritize receiving comprehensive training in the utilization of statistics. When the finding of questionnaire was considered in that most of them stated receiving either adequate or advanced level of training in “using statistics” and majority of them highlighted the need for advanced or moderate level of training in the same aspect, it was evident that they attach great significance to this aspect. This could be stemming from some factors including research and data analysis, data-driven teaching, professional development and growth and decision-making. As discussed previously, EFL teachers often engage in educational research or conduct assessments and evaluations. Training in statistics equips them with the skills to analyze and interpret data effectively, enabling them to make evidence-based decisions. It allows teachers to conduct meaningful research studies, analyze survey results, and evaluate the effectiveness of instructional strategies or materials. As well as this, having statistical knowledge empowers EFL teachers to analyze student performance

data and identify patterns or trends. They can use statistical techniques to examine student strengths and weaknesses, inform instructional decisions, and differentiate instruction based on individual or group needs. By analyzing data, teachers can adjust their teaching strategies, identify areas for intervention, and track student progress effectively. By having a solid understanding of statistics, EFL teachers can engage critically with research findings, evaluate the validity and generalizability of research studies, and stay updated with current trends and best practices in language teaching. This enhances their professional development and enables them to make evidence-based decisions in their classrooms. With statistical training, EFL teachers can also critically evaluate research studies, educational policies, or instructional interventions. They can make informed decisions based on evidence, ensuring that their teaching methods and strategies are grounded in research and best practices. Overall, getting training in statistics for in-service EFL teachers is important owing to the fact that it equips them with the skills to conduct research, analyze data, evaluate assessments, make data-driven instructional decisions and practice evidence-based decision-making. By having a strong statistical foundation, EFL teachers may desire to enhance their teaching practice, contribute to educational research, and improve student outcomes in language learning.

As for the second commonly mentioned topic, European Language Portfolio is also prioritized similar to the finding of questionnaire. As the finding of the questionnaire revealed, most of them stated sufficient or little training was offered to them and a similar percentage of them sought to take moderate to advanced training, which was also shown in the comments of interviewees. As can be seen in the explanations of interviewees, some of them explained the need to learn more about such topics as ELP when they come across them on social media or the internet. For instance, they wished to learn how it works to incorporate it for assessment purposes and the functionality. One of them stated she encounters “can-do” statements at the end of the units and wants to explore more believing that they worth exploring. Given the pivotal role of European Language Portfolio for all

language teachers, in-service English as a Foreign Language (EFL) teachers may ask for additional training in using the European Language Portfolio (ELP) for several reasons. The ELP is a valuable tool for language learners to document and reflect upon their language learning progress, and when teachers are trained in using it effectively, they can enhance their teaching practice and support their students more effectively. Several reasons could have affected the decision why in-service EFL teachers may need more training in using the ELP. Firstly, training provides teachers with a deeper understanding of the ELP's purpose, structure, and potential applications. It ensures that teachers are familiar with the various components of the portfolio, such as the language passport, the language biography, and the dossier, and how to guide students in using them. Secondly, it helps teachers integrate the ELP seamlessly into their teaching practice. They learn how to align the ELP with their curriculum, lesson planning, and assessment practices, ensuring that the portfolio becomes an integral part of the learning process rather than an add-on activity. As well as these, the ELP encourages learner autonomy by involving students in self-assessment, goal setting, and reflection. Training helps teachers guide students through these processes, teaching them how to set realistic goals, assess their own progress, and reflect on their strengths and weaknesses. This fosters a sense of ownership and responsibility for learning, empowering students to take an active role in their language development. With the ELP, teachers can better tailor instruction to meet the diverse needs of their students. Training equips teachers with the knowledge and skills to identify students' strengths, weaknesses, and learning preferences through portfolio assessment. This information helps them provide targeted feedback, offer appropriate resources, and personalize instruction to address individual student needs effectively. Furthermore, the ELP can boost learner motivation by providing a clear visual representation of progress and achievement. Training enables teachers to use the portfolio as a motivational tool, helping students see their language learning journey in a positive light. Teachers can guide students to celebrate their accomplishments and set new goals, fostering a sense of achievement and motivation to continue learning. Besides, the ELP encourages collaboration and

communication among teachers, learners, and parents/guardians. Through training, teachers learn how to engage in meaningful discussions with students and their families, sharing progress reports, setting goals, and exchanging feedback. This promotes a supportive learning environment and strengthens the teacher-student and teacher-parent relationships. In short, these teachers possibly consider that training in using the ELP is crucial as it enhances their ability to integrate the portfolio effectively into their teaching practice, promotes learner autonomy, facilitates individualized instruction, motivates learners, and fosters collaboration and communication. Thanks to the ELP, teachers can support their students' language learning journeys more effectively and promote lifelong learning habits.

The other most frequently stated topic was testing grammar and the participants expressed their wish to receive extra training. The finding of the questionnaire regarding this specific topic showed that most of the in-service language teachers stated having little to moderate level of training and similar percentage of them also expressed the need for training moderate to basic level. When both findings were interpreted along with that of interviews, it could be reported that they did not find their training in "testing grammar" completely sufficient for themselves and they asked for more training in this topic. Several factors could have led them to ask for extra training in this area. First of all, national examinations could be an important factor in that Turkey has national language proficiency exams, such as the Foreign Language Exam (YDS) and the Transition to Higher Education Exam (YKS) and these exams often include a significant grammar component. In-service language teachers need training in testing grammar to prepare their students effectively for these exams. Understanding the specific grammar constructs and the assessment methods used in these national exams is essential to provide appropriate instruction and practice. Secondly, curriculum requirements could also be effective in their need for additional training. In Turkey, language education is guided by a curriculum that specifies the grammar skills and knowledge students should acquire at each level. Language teachers need

training in testing grammar to align their assessments with the curriculum requirements. This ensures that the tests effectively evaluate students' understanding and application of grammar rules, allowing for accurate monitoring of progress and curriculum implementation. Regarding grammar' being focal point of curriculum, one of the interviewees emphasized that despite their efforts to avoid it, their teaching predominantly focuses on grammar, which is then evaluated through tests that typically follow traditional formats. She emphasized the importance of going beyond these traditional methods and acquiring knowledge about innovative approaches to assess grammar. Thirdly, students take the Undergraduate Placement Examination to enter universities, and this exam includes a foreign language section with a focus on grammar. In-service language teachers may need training in testing grammar to prepare their students adequately for this exam, ensuring that they develop the necessary grammar skills and are familiar with the exam format and requirements. Also, there is an emphasis on achieving certain language proficiency standards, such as the Common European Framework of Reference for Languages (CEFR). Testing grammar is crucial for assessing students' progress in meeting these standards. In-service language teachers need training to design tests that align with the CEFR levels, accurately measure grammar proficiency, and provide valid and reliable results. Besides, they could ask for training to equip themselves with the knowledge and skills to design meaningful grammar assessments and provide targeted feedback to students. This, in turn, improves their overall teaching effectiveness. By accurately assessing grammar proficiency and identifying specific areas for improvement, teachers can adapt their instructional strategies and support students' language development more effectively. The other point could be continuous professional development, which is essential for language teachers to enhance their pedagogical skills and stay updated with the latest trends in language education. They could have considered that it provides them with an opportunity to expand their expertise and develop a deeper understanding of assessment practices. This professional development contributes to their overall professional growth and promotes better teaching practices in the classroom. Shortly, it is highly possible that these teachers need more training in testing

grammar due to curriculum requirements, national and university entrance examinations, language proficiency standards, the need for enhanced teaching effectiveness, and the importance of continuous professional development. Through improving their assessment practices, they can ensure that students' grammar skills are effectively evaluated, leading to better language instruction and improved learning outcomes.

As well as these topics, other topics like testing listening, speaking and vocabulary were also brought up. Regarding testing listening, the questionnaire results showed that around 35% of them find the training sufficient and 31% of them regarded training as little. As for their extra training needs, a similar percentage of them (35%) ask for advanced training and 29% of them wished for intermediate training. In the same vein, in the interviews they requested more instruction for testing listening skill. It could be attributed to some issues like ensuring enhanced assessment accuracy, validity, authenticity, professional development, and diagnostic purpose. It is possible that these teachers want to take more training to develop the necessary skills to design and administer listening tests accurately. This ensures that students' listening abilities are evaluated reliably, providing a more accurate reflection of their true proficiency levels. Besides, extra training can help to align their listening tests with language learning objectives and curriculum standards. This ensures that the test content, format, and scoring criteria are appropriate, resulting in valid assessments that measure the intended listening skills effectively. It also empowers teachers to create listening assessments that mirror real-life language situations. Authentic listening tasks provide students with opportunities to practice and develop skills necessary for understanding spoken language in various contexts, thereby enhancing their overall communicative competence. Some of them mentioned that as English is a dominant language, and their students watch English movies without grasping most of the conversation. When students asked them for ways to develop their skills, they mostly recommended listening to audios every day for a certain time period to enhance their skills. Moreover, they are conscious of the fact that they should do something more authentic

instead of just exposing them to certain audios artificially. As well as this, they figured out there should be methods that are available for assessing their progress in the essential listening skill required for everyday situations. The other participant pointed out that they just know open something as listening activity and requiring learners to either fill in the blanks or choose the best or correct answer among options, which he regarded as an insufficient practice. He asked for to be provided with theory-based course. As seen in these, they may ask for more training for authenticity purposes and diversifying the ways to assess students. As in other constructs, ongoing training in "testing listening" ensures that teachers stay updated with the latest research, methodologies, and best practices. This continuous professional development enables them to adapt their teaching practices and assessment strategies to meet the evolving needs of their students effectively. The other reason can lie in the diagnostic insights. They may wish for further instruction as adequate training equips teachers with the skills to analyze and interpret students' performance on listening tests accurately. This allows them to identify individual strengths and weaknesses, enabling tailored instruction to address specific needs and improve students' listening proficiency. Overall, additional training in "testing listening" for in-service EFL teachers is crucial and this improves the quality of language instruction and contributes to students' language learning progress.

Testing speaking was another topic that was highlighted for further training by interviewees. When the findings of the questionnaire were reconsidered, it was seen that in-service language teachers did not find training completely sufficient as nearly one third of them stated the amount of training received as little. As for their level of further training needs, it was found that the majority of them requested moderate to advanced instruction, which was also in line with the findings of the interviews. The common thing in the interviewees' responses was that for something to be sufficient, there should be more than 1 course. Especially, one participant stated that how it was possible to cover all topics in assessment only in 1 course. This is also an important point to reconsider given that they

find the number of assessment courses in their education as limited. In addition to this, several reasons can lie in their desire for more training. It could be related to the challenging nature of assessing speaking skill as it involves evaluating the learner's ability to communicate effectively in real-time. Teachers need training to develop valid and reliable assessment techniques to ensure that they accurately evaluate students' speaking abilities. Without proper training, teachers may struggle to design appropriate assessments and may inadvertently introduce biases or inaccurately gauge students' proficiency levels. In addition, designing speaking tasks that effectively measure a student's language skills requires expertise and training. Teachers need to create tasks that encourage learners to demonstrate their ability to produce accurate and meaningful language in various contexts. Training helps teachers understand how to create tasks that are aligned with learning objectives, promote authentic communication, and elicit different language functions and structures. Moreover, they may wish for further training as it enables them to offer specific, targeted feedback that focuses on areas for improvement while also acknowledging students' strengths. It equips teachers with strategies for delivering feedback in a supportive and encouraging manner, ensuring that learners understand how to enhance their speaking abilities. Apart from these, as each student has unique strengths and weaknesses along with learning styles, receiving training can equip them with strategies for assessing individual learner needs and tailoring speaking assessments accordingly. It helps them identify areas where students require additional support, such as pronunciation, vocabulary, or grammar, and design interventions to address these specific needs effectively. The other benefit could be that such training empowers teachers to design activities and exercises that encourage students to practice speaking outside of formal assessment contexts. Such practice opportunities foster fluency, accuracy, and confidence in learners. By incorporating speaking practice into their instructional plans, teachers can better prepare students for real-world language use. Shortly, it is important to provide in-service language teachers with training in testing speaking because it enhances their ability to evaluate and develop students' speaking skills accurately and these teachers are conscious of these realities.

Effective assessment and feedback processes support learner motivation, guide instructional planning, and contribute to overall language proficiency development. With appropriate training, teachers can confidently assess speaking skills, provide meaningful feedback, and create a supportive learning environment that fosters successful language acquisition.

The other topic was “testing vocabulary” and these participants asked to be provided with further training as they did not find it sufficient in their undergraduate training. Similarly, the results of the questionnaire showed that most of them found training that was offered to them not totally enough for them, and they wished to receive more training in that many of them stated further training especially at moderate to advanced training levels. This could be related to many factors and one of them could be applying vocabulary assessment accurately. They could be stated to be aware of the fact that vocabulary knowledge is a crucial component of language proficiency. Assessing vocabulary effectively requires teachers to have a deep understanding of different aspects of vocabulary, including breadth, depth, and usage. Training helps teachers develop appropriate assessment techniques that evaluate learners' vocabulary acquisition accurately. It enables them to design tasks that assess learners' ability to understand and use a wide range of vocabulary items in various contexts. Also, designing vocabulary tasks that align with learning objectives and accurately measure learners' vocabulary proficiency requires expertise. Teachers need training to create tasks that assess different aspects of vocabulary, such as word meaning, collocations, word formation, and usage. As they stated, it is not just about requiring students to say the meaning of the target words in their native language. They wished to find alternative and creative ways to assess vocabulary knowledge. They also need to ensure that tasks provide learners with opportunities to demonstrate their ability to use vocabulary in meaningful ways. Training equips them with the skills to design varied and engaging vocabulary assessment tasks. The other reason could be differentiating vocabulary levels. Learners often have different levels of vocabulary proficiency. Training

in testing vocabulary helps teachers differentiate vocabulary assessments to cater to learners' individual needs. They learn how to design tasks that challenge and stretch more advanced learners while providing appropriate support for those at lower proficiency levels. This differentiation ensures that learners are assessed at their appropriate vocabulary level and allows for targeted vocabulary instruction. After assessing vocabulary performance, teachers need to provide meaningful feedback that supports learners in expanding their vocabulary knowledge. Training helps teachers develop strategies for giving specific and constructive feedback on vocabulary use and encourages learners to actively engage with feedback. Effective feedback contributes to vocabulary growth and enables learners to make informed decisions about further vocabulary development. Moreover, this training could inform instructional planning. Since vocabulary assessment outcomes provide valuable insights into learners' strengths and weaknesses, training in testing vocabulary equips teachers with the knowledge to analyze assessment results effectively and use them to inform instructional planning. They can identify areas where learners require additional support or specific vocabulary instruction, enabling them to design targeted interventions and provide tailored vocabulary instruction to meet learners' needs. It is crystal clear that vocabulary is essential for effective language use, including reading, writing, listening, and speaking. By assessing learners' vocabulary proficiency, teachers can identify areas that need improvement and design activities and exercises to promote vocabulary development across various language skills. Training in testing vocabulary enables teachers to integrate vocabulary assessment into a comprehensive language teaching and learning approach. In summary, providing in-service language teachers with training in testing vocabulary is crucial because it enhances their ability to assess vocabulary knowledge accurately, design appropriate vocabulary tasks, provide effective feedback, and inform instructional planning. This training empowers teachers to support learners' vocabulary development and facilitate their overall language proficiency growth.

When these topics are analyzed, it was seen that both language skills and knowledge are put emphasis significantly and they want to explore these more and to find alternative and more engaging ways to assess these subjects as they regard themselves as not adequate for the amount of knowledge they have and ways to assess these. This could be related to the fact that language learning primarily focuses on developing communication skills and linguistic competence and assessing language skills and knowledge is directly tied to this core objective. It measures students' ability to effectively communicate, understand, and interact using the target language. Especially their focus on testing listening and speaking along with grammar and vocabulary could be related to the greater focus on communicating, understanding and interacting in the target language. Getting training in these skills and knowledge ensures that assessments align with the primary goal of language education, which could be underlying reasons for them to prioritize these fields. The other reason could be that language skills and knowledge are interconnected. Testing language skills often requires assessing students' underlying knowledge of grammar, vocabulary, and other language components. Training teachers in testing language skills and knowledge allows them to understand the interplay between these elements and create assessments that capture their integration. They could look for specialized expertise in that language skills and knowledge testing requires specific expertise and understanding of the complexities of language acquisition. Language teachers need to possess in-depth knowledge of linguistic principles, language development, and the various components of language proficiency. They may wish to equip themselves with the necessary skills to assess language abilities accurately. The multidimensionality of language skills can be also influential in their wish for extra training. Assessing language skills encompasses multiple dimensions, including fluency, accuracy, vocabulary usage, grammatical structures, pronunciation, and discourse coherence. They could have considered that receiving training will equip them with the knowledge to design assessments that capture the complexity of these dimensions and provide comprehensive feedback to students. All in all, the emphasis on language skills and knowledge testing for

language teachers stems from the unique nature of language acquisition and the central role language proficiency plays in language education.

Surprisingly, these participants focused more on testing listening and speaking skills compared to the other two skills, which are reading and writing. The importance of training in testing listening and speaking skills, compared to reading and writing skills, for language teachers is not necessarily an issue of one being more important than the other. Both sets of skills are vital in language learning, and teachers should ideally receive training in assessing all four language skills comprehensively. However, there could be a few reasons why training in testing listening and speaking skills may be emphasized more. Firstly, listening and speaking skills are essential for interactive communication. In real-life language use, individuals engage in conversations, discussions, and negotiations where effective listening and speaking skills are crucial. Training in these skills helps them evaluate students' ability to comprehend spoken language, respond appropriately, and engage in meaningful conversations. Secondly, speaking is an active skill that requires students to produce language spontaneously. It involves combining vocabulary, grammar, pronunciation, and discourse skills in real-time. Training in assessing speaking skills helps them design tasks that elicit authentic responses and evaluate students' ability to express themselves fluently, accurately, and coherently. Furthermore, listening assessments often involve authentic audio materials such as recordings of conversations, lectures, or news broadcasts. These materials expose students to natural language use, various accents, and different speech rates. Further training in assessing listening skills helps them select and create appropriate listening tasks that reflect real-world language contexts and develop students' listening comprehension abilities. Also, speaking assessments provide opportunities for students to practice and demonstrate their language skills, contributing to their confidence and motivation. When students receive constructive feedback on their speaking abilities, it boosts their self-esteem and encourages further development. They could seek to receive more training to help them to create a positive learning environment

that nurtures students' confidence in their language abilities. On the other hand, many language education frameworks and proficiency scales, such as the Common European Framework of Reference for Languages (CEFR), emphasize the importance of oral proficiency. These frameworks recognize the significance of oral communication skills in language learning and highlight the need for training teachers to assess and evaluate speaking abilities accordingly. Overall, while training in assessing listening and speaking skills may be emphasized, it is important to recognize that reading and writing skills are equally important in language education. These skills play a crucial role in comprehension, written communication, and academic success. Language teachers ideally require training in testing all four skills to provide a well-rounded assessment and support the development of students' overall language proficiency.

The next interview question was related to whether these in-service EFL language teachers considered it is crucial for them to possess expertise in LTA to enhance the overall quality of learning and teaching. When their responses were analyzed, 13 out of 20 expressed a strong belief in the high importance of possessing expertise in LTA, providing justifications to support their viewpoint. On the other hand, four participants considered it partially important, and three participants expressed that they did not view it as crucial in any way. To start with justifications of those who expressed their strong belief in the crucial role of expertise in LTA, they emphasize the interdependence of teaching, testing, and learning and firmly believe in the significance of having competence in Language Testing and Assessment (LTA). One of them illustrated this connection by highlighting the time dedicated to test practices in comparison to teaching, thus elucidating the relationship between these processes. Similarly, the other one supported his belief by considering a holistic perspective on testing, teaching, and learning. They viewed testing as a means to identify and enhance instruction and learning. As seen in their explanations, they are informed about the relationship between teaching, learning, and assessment, which is integral to the educational process. Firstly, they are clearly aware of the fact that

assessment informs teaching by providing valuable feedback on students' progress and understanding. It helps teachers identify areas of strength and weakness, allowing them to tailor their instructional strategies to meet students' individual needs. Assessment data guides instructional decisions, such as adjusting teaching methods, pacing, and content. Secondly, assessment serves as a guide for student learning. It provides students with feedback on their performance, highlighting areas that require improvement. This feedback helps students reflect on their progress, identify their strengths and weaknesses, and make necessary adjustments in their learning strategies. They consider that assessment supports the development of metacognitive skills and promotes a deeper understanding of the subject matter. They could also have figured out that assessment plays a crucial role in setting clear learning goals and objectives. It helps educators identify the knowledge, skills, and competencies that students need to acquire. Assessments are aligned with these goals and provide benchmarks to measure students' progress towards achieving them. This alignment ensures that teaching and learning activities are focused and purposeful. The other relation is related to monitoring progress as assessment allows for the ongoing monitoring of student progress. Regular assessments provide insights into students' learning trajectories and help identify any gaps or areas where additional support may be needed. By monitoring progress, teachers can intervene promptly, provide targeted assistance, and ensure that students stay on track towards meeting their learning goals. The other factor could be accountability in that assessment promotes accountability for both teachers and students. It holds teachers accountable for the effectiveness of their instruction by providing evidence of student learning outcomes. Similarly, students are held accountable for their own learning as assessments measure their understanding and performance. Assessment fosters a sense of responsibility and encourages students to take an active role in their educational journey. Also, it can enhance student motivation by providing opportunities for recognition and achievement. Well-designed assessments provide a sense of accomplishment when students can demonstrate their learning and progress. This positive reinforcement encourages students to stay engaged, strive for

improvement, and take ownership of their learning. Along with these, professional development could be influential in their viewpoint, too. Assessment expertise is an essential component of a language teacher's professional growth. Teachers with assessment expertise are more likely to engage in ongoing professional development, keeping abreast of current assessment practices and research. This continuous learning helps teachers refine their teaching strategies, enhance their assessment methods, and stay informed about the latest trends and advancements in language assessment. Overall, teaching, learning, and assessment are interconnected components of the educational process. Effective assessment practices inform instruction, guide learning, set goals, monitor progress, enhance motivation, and promote accountability. When these three elements work together harmoniously, they create an optimal learning environment that supports student growth and achievement.

Others explained the significance of competency by mentioning the importance of employing appropriate approaches to testing and assessment, staying up-to-date, particularly in terms of assessment methods as they believed in exposing students to modern testing practices. Others also highlighted this aspect of competency by drawing comparisons between traditional methods, such as filling in the gaps and more contemporary approaches such as the use of portfolios. Assessment competency is emphasized by them as it is crucial for language teachers by allowing for accurate language proficiency evaluation, differentiated instruction, targeted feedback, curriculum design and adaptation, monitoring language development, and meeting accountability standards. It supports effective language teaching and ensures that students receive the necessary support to achieve their language learning goals. Staying abreast with contemporary assessment methods is important for language teachers as it allows them to reflect current language standards, meet changing educational needs, enhance assessment validity, promote learner engagement, leverage technological advances, and foster professional

growth. By utilizing these methods, teachers can provide effective and relevant language assessments that support student learning and success.

They have also handled the subject by pointing out the various education stakeholders in terms of providing them with accurate information about learners' success or weak points to be developed. It allows stakeholders to see the measures taken to evaluate student performance and the criteria used for assessment. This fosters trust and confidence in the teaching and learning process. Unless they provide such information to stakeholders, they can't promote accountability and transparency in the educational process.

Regarding the remaining teachers, four of them acknowledged the importance of testing competence, albeit to a certain degree. What is common in their remarks is that they are responsible for providing students with guidance and feedback on their progress and areas for improvement. They stated that one of their tasks is to make students informed about their progress, which is not only through announcing exam scores but also through guiding them. They seem to disregard the other benefits of competency in LTA, which were itemized above.

Conversely, three of the in-service EFL teachers did not perceive LTA competence as important and viewed testing as distinct. The common aspect among these three responses was that they did not connect the quality of teaching and learning processes to testing expertise. In fact, one of them explicitly emphasized a clear distinction between the two by pointing out that there is not any relation between testing and teaching. Another surprising comment was that one of them stated the field of testing as another separate field of expertise and it was unreasonable to expect them to be experts in testing as responsibility falls upon other professionals dedicated to testing worldwide. The other one mentioned that their task was to teach the subject in a timely manner and keep abreast of the latest teaching methods, which reveals that they distinguish teaching and assessment from each other. Shortly, they did not believe in the idea that expertise in LTA is integral for

enhancing teaching and learning. There could be underlying reasons for this. In the first place, some teachers may not be fully aware of the impact that LTA expertise can have on teaching and learning outcomes. They may not have received adequate training or exposure to the benefits of incorporating effective assessment practices into their teaching. In the second place, they may not have received adequate training or support in LTA during their professional development. Without the necessary knowledge and skills, they may not recognize the importance of LTA expertise for enhancing teaching and learning outcomes. Moreover, teaching involves numerous responsibilities, including planning lessons, delivering content, and providing feedback to students. Teachers may perceive LTA expertise as time-consuming, requiring additional effort to design and implement assessments, analyze data, and provide meaningful feedback. This can lead to a lack of emphasis on developing LTA competency. If they were exposed to more training in LTA and informed about washback effect focusing on the impact of testing on teaching learning and curriculum, their awareness could increase regarding the benefits of including efficient assessment practices into their teaching. Addressing these factors through training, awareness-building, resource provision, and ongoing support can help them understand the benefits of LTA competency and its potential impact on teaching and learning quality.

In the tenth interview question, they were asked to indicate what types of challenges and obstacles were encountered in their assessment and testing practices as language teachers and how they managed to address and overcome these challenges. Half of the participants mentioned having difficulty in “preparing tests”. The challenging aspect of test preparation was expressed by them as it involves the need to focus on multiple aspects simultaneously. Specifically, some of them said that they prefer adapting ready-made tests rather than creating their own tests as they are not experts at designing tests and they do not have much technical knowledge. They have some dilemmas in what to give their attention to like allocation and varieties of question types. To tackle this problem, they sometimes make use of ready tests and just adapt when required. To overcome these

problems, professional development opportunities focused on test preparation, access to reliable assessment resources, and collaboration with colleagues can be beneficial. Additionally, seeking guidance from assessment experts or employing established assessment tools and frameworks can support teachers in creating effective and reliable tests.

What others mentioned as difficulty was reliability and many of them highlighted that they could not check and ensure everything in the classroom including students, physical environment like sound system and computer. They encounter uncertainties regarding how to handle the exam or exam results. While they were employing tests in the classrooms, they experienced reliability-related issues. One of them is student-related factors. Students may experience test anxiety, lack of motivation, or unfamiliarity with the assessment format, which can affect the reliability of their performance. To address these issues, teachers can create a supportive classroom environment, provide test-taking strategies, and offer practice opportunities to familiarize students with the assessment format. Another problem could be the sound system or technical issues. Problems with the sound system or technical equipment like computers or photocopiers can lead to inconsistencies or disruptions during the assessment process. To mitigate these challenges, teachers should ensure that all equipment is in proper working condition before the assessment. It is also helpful to have backup plans in case of technical difficulties and to communicate clearly with students about any changes or adjustments. Also, standardization and scoring could be problematic for them. Ensuring consistent scoring practices among different teachers or examiners can be a challenge. To enhance reliability, it is important to establish clear scoring criteria and guidelines, provide training to teachers/examiners on scoring procedures, and conduct regular calibration meetings or workshops to maintain consistency in scoring. Furthermore, test administration procedures could be troublesome. Inconsistent or unclear test administration procedures can impact the reliability of assessments. Teachers should provide clear instructions to students regarding test expectations, time limits, and any

specific guidelines. It is also crucial to monitor the test administration process to address any issues or deviations that may arise. Regularly reviewing and reflecting on assessment practices can help identify and address reliability issues. Teachers can analyze student performance data, seek input from colleagues, and engage in professional development activities focused on assessment and reliability. This ongoing process of feedback and reflection can lead to continuous improvement in assessment practices. Overall, addressing reliability issues requires a combination of clear communication, appropriate training, consistent procedures, and ongoing reflection and improvement. By implementing these strategies, in-service EFL teachers can enhance the reliability of their assessments and promote fair and accurate evaluation of student performance.

Some expressed concerns and challenges about validity highlighting issues related to the content and construct validity though they did not explicitly use the terms of “content validity” or “construct validity”. According to them, ensuring comprehensive coverage of all the topics and desired language proficiency that they aimed to gauge in assessments is challenging. Teachers may have a limited understanding of validity concepts and how they apply to language assessments. To address this, professional development programs or workshops focused on assessment literacy and validity can be beneficial. Providing teachers with resources and guidance on validity can enhance their understanding and application of these concepts. Besides, ensuring that assessments accurately measure the intended language constructs or content can be difficult. Teachers should carefully align their assessments with the learning objectives of the curriculum or course. Regularly reviewing and revising assessments in collaboration with colleagues can help ensure that they accurately reflect the desired language constructs and content. They might also struggle with developing assessment items that effectively measure the intended constructs. Collaborating with assessment experts or colleagues who have experience in assessment design can provide valuable insights. Teachers can also utilize existing assessment resources or frameworks to guide the development of valid assessment items.

On the other hand, validity should be an ongoing consideration in the assessment process. Teachers should regularly evaluate the validity of their assessments by examining the relationship between the test scores and the desired language constructs. Conducting item analysis, seeking student feedback, and engaging in peer review can assist in identifying and addressing potential validity issues. Collaborating with other EFL teachers and assessment experts can help address validity concerns. Sharing best practices, discussing assessment strategies, and seeking feedback from colleagues can enhance teachers' ability to design valid assessments. Creating professional learning communities or participating in assessment-focused forums can foster collaboration and support in addressing validity challenges. By addressing these issues through professional development, alignment with learning objectives, collaboration, and ongoing evaluation, in-service EFL teachers can enhance the validity of their assessments and ensure that they accurately measure the desired language constructs and content.

Others mentioned difficulty in providing feedback, citing the importance of considering learners' motivation and well-being. Specifically, one of them highlighted the potential risk of learners' taking feedback personally and emphasized the need to carefully consider this issue to prevent such situations from occurring. In-service EFL teachers may encounter challenges in giving feedback, especially when dealing with sensitive students who may take comments personally. Some students may be more sensitive to feedback, perceiving it as criticism or personal attacks. This can make it challenging for teachers to provide constructive feedback without negatively impacting students' motivation and self-esteem. To solve this, teachers can employ strategies such as using a supportive and encouraging tone when giving feedback, focusing on specific strengths alongside areas for improvement, and emphasizing the developmental nature of feedback. Creating a safe and respectful classroom environment where students feel comfortable expressing their concerns can also help address sensitivity issues. Students may have a tendency to take feedback personally, interpreting it as a reflection of their worth or abilities. This can hinder

their willingness to accept and learn from feedback. To tackle this, teachers can help students develop a growth mindset, emphasizing that feedback is an opportunity for improvement and not a judgment of their worth. Encouraging self-reflection and self-assessment can also empower students to take ownership of their learning and view feedback as a valuable tool for their progress. Moreover, inadequate communication or unclear guidance on the purpose and process of feedback can contribute to misunderstandings and students taking comments personally. To solve this, they should clearly communicate the objectives and intended benefits of feedback to students. They can provide explicit instructions on how to interpret and use feedback constructively. Encouraging students to ask questions and seek clarification can help in fostering a better understanding of feedback. Sometimes, it could be challenging for teachers to provide individualized feedback to a large number of students within limited time constraints, leading to generic or less personalized feedback. For solving this, teachers can explore strategies such as using rubrics or checklists to provide specific feedback tailored to individual students' needs. Grouping students with similar areas for improvement can also facilitate more targeted feedback. Additionally, leveraging technology can enable more efficient and personalized feedback, such as audio or video recordings or online platforms that allow for individualized comments. Teachers may benefit from training and professional development opportunities focused on effective feedback practices, including strategies for providing constructive feedback and managing sensitive situations. Engaging in workshops, seminars, or professional learning communities that address feedback techniques and best practices can enhance teachers' skills and confidence in delivering effective feedback. Collaborating with colleagues and sharing experiences can also contribute to professional growth in this area. By implementing these strategies, in-service EFL teachers can navigate the challenges associated with giving feedback, promote a positive learning environment, and support students in effectively utilizing feedback for their language development and growth.

When their solutions for solving such problems were reanalyzed, it was seen that they seek advice from colleagues, conduct research, read relevant materials online, engage in research and explore relevant forums or newsgroups to find potential solutions for their problems. Some teachers prefer discussing their problems with colleagues due to their familiarity with the context, students, and school, as well as the sense of security, confidence, and knowledge of strengths and preferences. There could be some factors underlying their preferences. Colleagues who work in the same context have firsthand knowledge of the challenges and specific circumstances teachers face. Consulting them can provide valuable insights and solutions tailored to the local context. Also, sharing problems with workmates allows for the exchange of experiences and perspectives. Other teachers may have encountered similar issues and can offer practical advice based on their own experiences and solutions they have implemented. Moreover, seeking guidance from workmates can provide emotional support and reassurance. Knowing that others have faced similar challenges and overcome them successfully can boost teachers' confidence in finding solutions. Colleagues also bring diverse skills, expertise, and teaching styles to the table. Consulting with them can provide a range of perspectives and approaches to problem-solving, expanding the possibilities for finding effective solutions. However, one expressed strong opposition to seeking advice from colleagues based on past experiences, which could be related to negative experience that he had. As for conducting research and exploring relevant materials online, they enable teachers to access a vast array of resources, scholarly articles, best practices, and discussion forums. These sources can offer insights, strategies, and alternative approaches to address specific challenges. Engaging in research and online learning keeps teachers updated on the latest developments, research findings, and innovative approaches in language teaching. This continuous professional development enhances their problem-solving skills and equips them with new tools and strategies. By consulting with workmates, conducting research, and accessing online resources, in-service EFL teachers can tap into collective knowledge,

gain new perspectives, and benefit from the expertise of others, ultimately improving their problem-solving abilities and enhancing their teaching practices.

Forums and newsgroups also provide a platform for teachers to connect with educators from different backgrounds and experiences. By exploring these online communities, teachers can gain access to a wide range of perspectives and ideas from professionals worldwide. This diversity of viewpoints can lead to innovative solutions that they may not have considered otherwise. In addition, they offer a space for teachers to share their own experiences and challenges. By engaging in discussions and reading about others' experiences, teachers can find validation, empathy, and insights into similar problems they are facing. This shared knowledge can be a valuable source of practical advice and strategies. Forums and newsgroups also foster a collaborative environment where teachers can collectively work together to solve problems. By posting questions or sharing their challenges, teachers can receive feedback, suggestions, and potential solutions from a community of educators. This collaborative problem-solving approach taps into the collective wisdom of the teaching community and can lead to more effective problem-solving outcomes. Upon exploring relevant forums or newsgroups, in-service EFL teachers can tap into a vast network of educators, benefit from shared experiences and knowledge, and find potential solutions to their problems or challenges in a collaborative and resourceful manner.

The others explained that encountering problems is a normal part of their profession, and if the issues are not significant, they might temporarily ignore them until finding a solution. However, one of them told that if the same problems persisted, she would make effort to devise essential solutions. It could stem from the factor that teachers prioritize their time and energy on issues that are deemed critical or have a significant impact on student learning outcomes. They may consider some problems or challenges in assessment to be less urgent or less likely to affect the overall assessment process or students' performance. Also, dealing with numerous problems or challenges simultaneously can be overwhelming

and lead to decision fatigue. Ignoring certain assessment-related problems temporarily may be a coping mechanism to manage workload and prioritize other pressing responsibilities.

In the last interview question, the interviewees were asked if they had any additional comments or inquiries, and only two of them posed questions in response.

One of the interviewees had a prompt question and expressed that when she considered the aspects they discussed during the questions and the survey she conducted, she had a strong desire to enhance her understanding of language testing. She asked the ways to further develop herself in this area and she asked for any suggestions on the best approach to achieve that. The other interviewee similarly added that her awareness is raised, and she started to consider possibilities to enhance her knowledge in the field. She also expressed that reading books is one of the ways, but she sought to have a more hands-on and engaging approach to accomplish this. From these statements, it could be stated that during the current study, they could be stated to find the opportunity to reflect on LTA field and their professional development especially in the field of language testing and assessment.

All in all, when the findings of interviews were considered as a whole within the scope of the training received and needed, it was found out that the training in their university education regarding LTA field was not found totally enough for most of them. In the same vein, many of them evaluated taking one testing and evaluation course inadequate for incorporating all constituents of LTA and they put forward many topics to explore more for their professional development and increasing their language assessment literacy as well as increasing the quality of teaching and learning processes by receiving more training. These findings were in harmony with those of the questionnaire. Some problems were expressed to be arising from insufficient training offered to them and they try to overcome through some ways like collaborating with workmates, conducting online research and remembering prior knowledge or experience. On the other hand, the majority

of them expressed strong belief in the competency in LTA for enhancing the quality of education.

Overall, findings of questionnaires and interviews reveal that in-service language teachers perceive the training they receive in Language Testing and Assessment (LTA) as not adequate totally. While they acknowledge the value of the training they have received, they express a strong desire for additional training, particularly in specific areas. This indicates that there are specific aspects within the field of LTA where teachers feel the need for more comprehensive knowledge and skill development. These results highlight the importance of ongoing professional development opportunities for language teachers, catering to their evolving needs and addressing the areas they identify as requiring further training. By offering targeted and tailored training programs, educational institutions and professional development providers can empower in-service language teachers to enhance their LTA expertise, thereby improving the quality of their assessment practices and ultimately benefiting their students' language learning journey. It can be argued that in-service English language teachers have a clear understanding of the benefits and significance of Language Testing and Assessment (LTA) courses for their Language Assessment Literacy (LAL) levels and overall proficiency. They recognize the valuable impact of LTA training on their professional growth and classroom practices, and they seek to address the gaps in their LAL through receiving further targeted instruction.

As put forward by Popham (2009), Siegel and Wissehr (2011), and Taylor (2009), the bulk of language teachers' knowledge in assessment primarily stems from the courses offered in their teacher education programs and their practical experience. Therefore, these findings point to the insufficiency of the number and content of the courses in LTA offered and further instruction need for in-service trainings was uttered explicitly as in the prior studies (Ballidag & Inan Karagul, 2021; Hasselgreen, Carlsen, & Helness, 2004; Salami & Alharthi, 2022; Vogt & Tsagari, 2014).

Chapter 5

Conclusion and Suggestions

Summary of the Study

Considering the increasing significance of evaluations and measurements in the field of language, it is crucial for language teachers to possess adequate knowledge and education in language assessment. Language assessment literacy (LAL) holds great importance in language education as it enables teachers to comprehend and apply information regarding students' performance, thus facilitating the language learning process. However, despite the pivotal role of assessment in language education, teachers often lack sufficient training in language testing and assessment, resulting in deficiencies in knowledge, practice, experience, and confidence in conducting assessments. Upon reviewing existing studies in this area in Turkey and the world, it becomes evident that limited research has been conducted in Turkey regarding the LAL levels and perceived training needs of in-service language teachers who are responsible for both classroom-based assessments and standardized language tests in their professional roles. Therefore, it becomes evident that further investigation is required to determine the LAL levels and training needs of English language teachers currently employed in diverse educational settings in Turkey.

The primary objective of this study is to explore the LAL levels and training needs of in-service English as a Foreign Language (EFL) teachers in order to address this significant gap in the field. Additionally, the study aims to examine whether the training needs of in-service English language teachers differ based on their level of experience and education. Furthermore, it seeks to investigate whether the training needs of in-service EFL teachers vary depending on the grade level they teach and the educational settings in which they work. Moreover, the study aims to gather English language teachers' accounts of their experiences with language testing and assessment practices within classroom contexts.

Overall, this study aims to analyze various dimensions, such as different educational settings, grade levels, levels of education, and experience, to examine the development of course-based LAL among in-service English language teachers, as well as their additional training needs and experiences in the field of language testing and assessment.

The researcher employed a sequential explanatory mixed method design, whereby quantitative data was collected and analyzed in the initial phase, followed by the collection and analysis of qualitative data in the subsequent phase of the study. This research design was utilized to provide an overall understanding of the LAL levels and requirements of in-service EFL teachers in relation to language testing and assessment. Additionally, it aimed to enhance this understanding by delving deeper into their training needs and experiences in language testing and assessment within the context of language classrooms. The quantitative part of the study involved the participation of 300 in-service EFL teachers. Among the participants, there were 171 female teachers and 129 male teachers, with ages ranging from 22 to 54 years old. In terms of teaching experience, the participants had varying levels of experience as English teachers. Specifically, 37.3% of them had 6-10 years of experience, 31.7% had 1-5 years of teaching experience, and the remaining 31% had 11 or more years of teaching experience. Moreover, 108 participants reported working in private schools whereas 99 and 93 participants were employed in public schools and private language schools, respectively. The participants also indicated working across different levels, including primary (n=102), secondary (n=101), and high school (n=97).

When asked about whether they had taken a testing and evaluation course specifically in English Language Teaching (ELT), 97% of the participants indicated that they had taken such a course, while only 3% stated that they had not. In terms of their educational backgrounds, 261 participants reported having obtained a Bachelor's degree, while twenty-seven participants mentioned holding or studying for Master's degrees, and 12 participants held PhD degrees. The majority of the participants were graduates of ELT departments while the remaining participants held Bachelor's degrees in other fields such

as English Linguistics, English Language and Literature, Linguistics, American Culture and Literature, and Translation and Interpreting. Among those with Master's degrees, over half of them studied ELT, while others had specialized in fields such as Curriculum and Instruction, English Language and Literature and Educational Administration. Regarding the PhD degrees, participants specialized in fields such as ELT, English Linguistics and English Language and Literature.

Additionally, the qualitative part of the study included the involvement of 20 volunteer in-service EFL teachers taking part in the quantitative part of the study. To ensure ethical considerations, pseudonyms were assigned to each interviewee. Among them, 6 hold a Master's degree. In terms of teaching experience, it ranges from 1-5 years to 11-15 years. Regarding the type of school they teach at, 6 participants work in private schools, 8 participants work in public schools, and 6 participants work in private language schools. Furthermore, their teaching levels vary, including primary school, secondary school and high school.

For the present study, a questionnaire with two parts was utilized to collect quantitative data from in-service English Language Teaching (ELT) teachers. The initial part of the questionnaire included inquiries about age, gender, education level, years of teaching experience, type of school/institution, and current grade level taught. The second part was adapted from Vogt and Tsagari's (2014) Teachers' Questionnaire, which was modified to consist of four sections in a Likert-scale format. These sections covered classroom-based language testing and assessment (LTA), purpose of testing, content and concepts of LTA, and knowledge of testing and assessment. Each section comprised two parts: Part A focused on the training received, while Part B addressed training needs. In Part A, in-service EFL teachers were asked to rate the extent of training they believed they had received during their undergraduate education using a 4-point Likert-type scale. In Part B, they were requested to rate their training requirements for the same items provided in the previous section, using a 4-point Likert-type scale.

Semi-structured interviews were conducted with in-service English as a Foreign Language (EFL) teachers employed in various educational settings, and these interviews were recorded for reference. The interview questions were adapted from Jeong's (2013) Instructor Survey and Instructor Interview Questions - Language Assessment Course. Some necessary modifications were made to align the questions with the objectives of the current study, including changes in wording and the addition of new questions. To ensure the validity of these questions, they were initially reviewed by the research supervisor, and the opinions of two teacher educators were sought. They were asked to assess the clarity and wording of the questions, which contributed to the clarity, content, and construct validity of the instrument while also identifying any redundancies (Brown, 2001). Subsequently, the instrument was pilot-tested with two in-service EFL teachers to refine its format and control its length. The interview comprised 11 open-ended questions, allowing the researcher to gather comprehensive data from the participants. These questions focused on the participants' educational background, their perceptions of the Testing and Evaluation in ELT course they received during their undergraduate education, including the adequacy and scope of the course. They also explored the implementation of Language Testing and Assessment practices, challenges faced by language teachers in testing and assessment, potential solutions to address these difficulties, and the training needs related to Language Testing and Assessment.

Prior to initiating the data collection process, approval was sought from the Hacettepe University Ethics Commission. Once the approval was obtained, in-service English as a Foreign Language (EFL) teachers from diverse formal and non-formal educational institutions, teaching at different grade levels, were invited to participate in the study through an online questionnaire. Participants were required to provide their informed consent by signing an official consent form. The researcher was available to answer any questions and provide clarifications as needed. Before administering the questionnaires to the participants, a pilot study was conducted with a group of in-service ELT teachers to

assess its effectiveness in an authentic context. This pilot study took place during the spring term of the 2020-2021 academic year. Participants were instructed to carefully read the instructions and questionnaire items, and to express any items they found unclear. Based on the feedback received from the pilot study, the instrument was revised accordingly. Subsequently, the main study was conducted during the fall term of the 2021-2022 academic year.

Regarding the interviews, a pilot study was conducted with two in-service EFL teachers prior to the actual interviews. This pilot study aimed to refine the format and length of the questions and allowed the researcher to assess their interviewing skills. It also helped identify any unclear aspects or issues related to the wording of the questions. The pilot study interviews were recorded to facilitate the enhancement of the interviewing skills. Following the successful pilot study, interviews were conducted with 20 in-service EFL teachers over the phone.

Quantitative data analysis in this study involved the use of descriptive and inferential statistics. To determine the appropriate inferential statistical test, a test of normality was conducted using the Kolmogorov-Smirnov and Shapiro-Wilk tests. The results indicated that the data followed a normal distribution, allowing for the application of parametric statistical tests. The first two research questions were addressed using descriptive statistics, such as percentages, frequencies, mean values, and standard deviation scores. These statistics were utilized to examine the training received and needed by in-service language teachers, by calculating mean values, frequencies, and percentages. For the third research question, one-way ANOVA was employed to explore whether there were statistically significant differences in training needs related to LTA among in-service EFL teachers based on factors such as the type of school/institution they teach at, the grade level they teach, their level of experience, and their education level. In cases where significant differences were found, a post hoc test (Tukey HSD) was used to further investigate the variables contributing to these differences, providing additional insights into the findings (Field, 2018).

The analysis of the qualitative data, corresponding to the last research question, was conducted using Grounded Theory (Strauss & Corbin, 1998). This involved the application of open and axial coding procedures. Open coding was the initial step, during which categories were identified. Axial coding focused on conceptual organization by establishing connections between categories and themes derived from the previous coding process. The codes were subsequently grouped to create broader categories and themes. After identifying the themes, a comparison was made between these themes and the questionnaire results to establish connections between the qualitative and quantitative data. This process aimed to validate the findings through triangulation, ensuring consistency and reliability in the overall conclusions.

The results of the study revealed that in-service EFL teachers considered their training in these areas during their university education to be somewhat insufficient, although not completely inadequate with an average score of 1.57 in all four domains. Among them, "content and concepts of LTA" received the highest average score ($M=1.73$, $SD=.42$). The second highest average score was for the domain of "knowledge of testing and assessment" ($M=1.56$, $SD=.53$), followed by the domain of "purposes of testing" ($M=1.51$, $SD=.65$). The domain of "classroom-focused LTA" had the lowest average score of 1.39 ($SD=.35$). In general, participants perceived that the training they received in the domain of "content and concepts of LTA" was better than in the other three domains, while the domain of "classroom-focused LTA" was rated lower compared to the others. Their LAL may be described as partially developed rather than fully developed as the amount of training is stated to be related to the development level of LAL (DeLuca & Johnson, 2017; Guerin, 2010; Lam, 2015; Mertler, 2004).

Furthermore, intermediate level of further training is expressed in these domains by the participants with the average score of 1.72. Among the domains of LTA, "content and concepts of LTA" received the highest average score ($M=1.94$, $SD=.384$). The second highest average score was for the domain of "purposes of testing" ($M=1.86$, $SD=.57$),

followed by the domain of "knowledge of testing and assessment" ($M=1.65$, $SD=.56$). The domain of "classroom-focused LTA" had the lowest average score of 1.43 ($SD=.386$). Overall, in-service EFL teachers expressed a greater need for further training in the domain of "content and concepts of LTA" compared to the other three domains, while the need for further training in "classroom-focused LTA" was perceived to be lower.

Another objective was to examine whether there were any notable differences in the training needs of participants based on their educational settings. The outcomes of the one-way ANOVA test indicated that there were no significant differences in the participants' further training needs in LTA in relation to their educational settings. Similarly, no significant differences were found among teachers working at different grade levels regarding their training needs in LTA. Additionally, the years of experience among teachers did not emerge as a significant factor in determining their future training needs in the field of EFL teaching. However, there were significant differences observed among teachers with different levels of education in terms of their further training needs in LTA ($F(2,297) = 3.542$, $p = .030$). This implies that the training needs of EFL teachers varied depending on their educational background. Specifically, the mean score for EFL teachers holding an MA degree ($M=1.85$, $SD=.28$) was significantly different from those with a BA degree ($M=1.70$, $SD=.26$) at a significance level of $p=.025$. However, no significant differences were found among the other groups of teachers in relation to their training needs and educational levels.

The findings of the qualitative data revealed that most of them expressed that Testing and Evaluation Course was inadequate for their needs. They identified several topics that they found most beneficial as language teachers. Analysis of the participants' listed topics revealed 8 components, including alternative assessment, validity, adapting prepared tests, practicality, informal/formal assessment, objective testing, formative/summative assessment, and testing reading. These topics were considered highly valuable in their actual classroom experiences. When discussing their beliefs regarding topics they would like to learn more about, the participants mentioned various

subjects such as statistics, alternative assessment, computer-based testing, reliability, approaches to language testing, giving feedback, testing integrated language skills, and validity. They shared that they often consult with their colleagues to decide on the assessment methods, considering it a practical and straightforward approach that emphasizes cooperation. They also mentioned utilizing the internet, referring to books, and drawing from their prior knowledge and experience as useful ways to choose appropriate assessment methods.

Regarding the adequacy of LTA training during their university education, most of them expressed that it was not satisfactory. When asked to list the areas in which they required further training, topics such as statistics, using ELP and testing grammar were mentioned more frequently. In terms of perceiving LTA competence as essential for enhancing their learning and teaching practices, 13 participants considered it highly important, providing justifications for their viewpoint. When asked to express the difficulties they encountered in LTA practices, participants highlighted areas such as reliability, test preparation and providing feedback.

Implications of the Study

This dissertation makes a valuable contribution to the existing body of literature concerning the LAL and training needs of in-service EFL teachers. Drawing on the study's findings and pertinent literature, the document enumerates the potential implications for pre-service teacher education and in-service teacher training.

Implications for Pre-service EFL Teacher Education

The proficiency of language teachers in language assessment (LTA) is considered crucial in second/foreign language education. Teacher education programs play a vital role in equipping language teachers with better competence in language assessment literacy (DeLuca & Klinger, 2010). When in-service language teachers express dissatisfaction with the training they receive in LTA, it highlights a significant problem that needs attention.

Consequently, the content of the courses offered in this program can be revised to address this issue. These courses can be enhanced to provide a sufficient level of training in both theoretical knowledge and practical application of assessment practices. Moreover, they can promote the language assessment literacy of pre-service language teachers and further develop the competence of in-service language teachers. To achieve this, the LTA course offered in the final year of the program could be revised to not only establish a solid foundation in language assessment but also incorporate the latest advancements in the field. This would enable language teachers to stay updated with innovative approaches and apply them effectively in their future assessment practices.

Another implication for undergraduate studies is related with the increasing demand for enhancing language testing and assessment (LTA) competence and language assessment literacy (LAL) along with its prominence for language teaching and learning. Currently, the undergraduate program includes only one LTA course, but it is necessary to incorporate more courses that provide opportunities for enhancing LAL and increasing their consciousness of the interconnectedness of teaching and learning and assessment. This interconnection could be ignored sometimes by them if they were not given explicit instruction. In the courses to be offered to prospective teachers, their understanding of this interconnection can be enhanced, which is quite critical for their future understanding and applications. Also, it has implications for classroom management in that LAL and LTA expertise also plays a role in effective classroom management. Pre-service EFL teachers who possess a high level of knowledge in these areas can create an environment that promotes language learning, communication, and engagement. They can design activities and assessments that encourage active participation and foster a supportive learning community. In addition, pre-service EFL teachers with strong LAL skills can seamlessly integrate language awareness into their instruction. They can explicitly teach language forms, functions, and features, which helps students to deepen their understanding and

application of the language. This fosters a more comprehensive and meaningful language learning experience.

Moreover, the needs of in-service English as a Foreign Language (EFL) teachers from various backgrounds regarding LTA training were expressed. Considering that the same line of thought can also be employed for pre-service teacher education in Turkey, it is recommended that changes be made to undergraduate programs to address this concern. This could involve the inclusion of additional courses specifically focused on LTA or the integration of practical demonstrations within existing courses to better equip future EFL teachers. All in all, it is recommended that pre-service teacher education programs offer elective courses or incorporate LTA components into the existing curriculum to meet this demand.

Additionally, it is recommended to arrange workshops that concentrate on language assessment. These workshops would provide a platform for experienced teachers to share their hands-on experiences, discussing the difficulties they encountered and offering suggestions for future practices within genuine teaching environments. Involving language testing and assessment experts in the workshops would offer guidance and create opportunities for participants to design language tests and assessments under their supervision. Instead of relying solely on theory-driven tests, pre-service teachers could engage in observatory practices. This would involve actively observing real language classes including the assessment carried out, taking notes, and reflecting on the experience. Moreover, undergraduate programs could incorporate more in-class practice of assessment, demonstrations, and hands-on experiences to better prepare pre-service teachers for their future roles. These approaches would deepen prospective language teachers' comprehension of LTA effectively and inspire them to engage in thoughtful reflections on real-world assessment practices.

Implications for In-service EFL Teacher Training

The field of English as a Foreign Language (EFL) teaching has been evolving rapidly, driven by the demands of globalization and the increasing importance of English as a lingua franca. In-service training plays a crucial role in ensuring that EFL teachers are equipped with the necessary knowledge and skills to effectively teach English to non-native speakers and assess their language knowledge and skills. Based on the result that in-service EFL teachers often find their education inadequate and express a need for intermediate level in-service training, this finding has significant pedagogical implications that need to be addressed in order to enhance their LAL and the quality of EFL instruction.

One major implication of the study is the importance of ongoing professional development for EFL teachers. In-service training is a means to update teachers' knowledge, refine their teaching strategies, and keep them abreast of the latest trends and methodologies in language testing and assessment. When teachers perceive their education as inadequate, it reflects a gap between their training and the evolving demands of the classroom. Therefore, it is crucial for educational institutions and policymakers to recognize the need for continuous training and provide EFL teachers with opportunities for professional growth.

Another pedagogical implication is the need for a targeted and tailored approach to in-service training. The study indicates that EFL teachers require intermediate level training, suggesting that a one-size-fits-all approach may not be effective. If training programs are not based on the needs of EFL teachers, their LAL and LTA expertise along with the satisfaction for the courses received may not enhance at the desired level. Therefore, future professional development programs should consider conducting needs-analysis studies to determine the needs and preferences of EFL teachers before implementation. In other words, in-service training programs should be designed to address specific needs identified by the teachers themselves, taking into account their prior education and experience. This requires conducting needs assessments and involving teachers in the planning and

development of training programs. By customizing the training to address teachers' specific areas of concern, institutions can ensure that the education provided is relevant and meaningful. Especially, the topics specified by the participants should be added to the content of further training programs though the other topics could also be incorporated into the course content for sound knowledge base.

Furthermore, the study highlights the importance of incorporating practical teaching components in in-service training programs. EFL teachers often face challenges in applying theoretical knowledge to their classrooms during their assessment practices. They need opportunities to practice and receive feedback on their assessment techniques and methods. Incorporating microteaching sessions, peer observations, and reflective practice can help bridge the gap between theory and practice. By engaging in hands-on activities, teachers can develop their pedagogical skills and gain confidence in their abilities, ultimately enhancing the quality of EFL instruction.

When the problems encountered in their assessment processes are considered, the identification of problems like reliability and test preparation has significant pedagogical implications for arranging in-service trainings accordingly. These findings highlight the need for targeted professional development programs that specifically address these challenges. In-service trainings should focus on enhancing teachers' understanding of assessment principles and practices, providing them with strategies and techniques to overcome such possible challenges that could occur in real-classroom settings. By arranging in-service trainings that directly address these identified issues, educational institutions can support EFL teachers in improving the quality and validity of their assessment practices, ultimately helping them in the problematic areas and enhancing the overall effectiveness of their instruction and promoting better learning outcomes for students.

The study also underscores the significance of technology integration in in-service training programs. The field of language testing has witnessed a rapid influx of digital tools and resources that can enhance language assessment experiences. In-service training

should equip EFL teachers with the knowledge and skills to effectively integrate technology into their assessment practices. This includes training on digital literacy, the use of educational apps and platforms and online assessment methods. By incorporating technology in training, teachers can enhance their pedagogical and LTA repertoire and adapt to the changing educational landscape.

The finding that in-service EFL teachers do not find their education adequate and express a need for intermediate level training carries important pedagogical implications. Continuous professional development, tailored in-service training programs, practical teaching components, and technology integration are key considerations to address the identified gaps. By investing in the ongoing growth and development of EFL teachers, educational institutions can ensure that teachers are equipped with the necessary skills and knowledge to provide high-quality assessment practices to learners. Ultimately, these pedagogical implications will contribute to improving EFL education and meeting the evolving needs of language teachers and learners in a globalized world.

On the other hand, the content of the further training could vary for different groups like BA, MA, and PhD based on their priorities and different contexts. At the BA level, it could be suggested that the focus of extra training in testing and assessment could be typically on building foundational knowledge and skills. The content may cover the fundamental principles of testing and assessment, including basic statistical concepts, test construction, test administration, and scoring procedures. Training may emphasize practical applications and techniques used in educational or workplace settings. B.A. level training often provides an introduction to various assessment tools and techniques without delving into advanced topics because of the limited number of testing and assessment courses. As for in-service training for individuals with a Master's degree, it typically builds upon the foundational knowledge acquired at the B.A. level. The content at this level may include more advanced topics, such as test validation procedures and statistical analysis in assessment. M.A. level training may also focus on the development and evaluation of

different types of assessments, including formative and summative assessments, performance assessments, and standardized tests. Participants may explore various approaches to test interpretation and reporting. Regarding the PhD level, extra training could include more specialized and advanced content. It may encompass complex topics like advanced statistical methods in testing and assessment, large-scale test development and analysis, differential item functioning and computerized adaptive testing. Ph.D. level training could emphasize research methodologies related to testing and assessment, including experimental design, validity research, reliability analysis, and test fairness studies. Participants may engage in critical discussions on emerging issues and current research in the field. Considering the topics listed by the present study's participants, they could be included in different education level training like PhD, MA and BA based on such consideration and needs. On the other hand, it's important to note that the exact content and curriculum of in-service training programs can vary between institutions and programs.

Limitations of the Study

In the data collection process, this study employed a convenience sampling technique. While this approach facilitated access to EFL teachers in various settings, it is important to recognize that not all EFL teachers in the target population had an equal opportunity to participate. Nevertheless, utilizing this method enabled the researcher to gather data from participants who met specific criteria and this study provides some information about the training needs of EFL teachers and LAL levels.

Additionally, it is worth noting that the current study focused exclusively on EFL teachers as participants while excluding other key stakeholders such as policy makers, program designers, school administrators, teacher trainers, Education faculty members, pre-service teachers, and language learners. Including the perspectives of these stakeholders would be valuable for informing in-service training programs. However, the

findings from this study can still be utilized to develop in-service training programs for LTA specifically tailored for EFL teachers in Turkey.

This study utilized a mixed-methods approach, employing both quantitative and qualitative research designs. Data were collected using questionnaires and interviews to comprehensively capture the level of language assessment literacy among teachers and identify their training needs in LTA. In addition to these, classroom observations or analysis of teachers' assessment samples could be included in the study. Incorporating multiple data sources can provide a more comprehensive understanding of their assessment practices.

Suggestions for Further Research

Taking into account the limitations of the study, the following recommendations are put forward to inform future research investigations:

Firstly, despite having a large number of participants, the current study does not encompass all EFL language teachers working in Turkey. It would have been advantageous to collect data from a broader range of EFL teachers in Turkey, which would have facilitated the generalization of the results to all in-service English language teachers in the country. Further studies may examine the LAL levels and in-service training needs of in-service EFL teachers by including a wider spectrum of EFL teachers in Turkey. On the other hand, EFL teachers from other countries or regions could also be included in the further studies and comparative analysis could be conducted in that multiple countries or regions could offer insights into the EFL teachers' training needs and LAL levels across different contexts. Comparisons could be made between countries with different educational systems, levels of resources, or cultural contexts, which may contribute to a more nuanced understanding of the factors influencing training needs and LAL levels.

Secondly, to address the limitation of excluding key stakeholders in the current study, future research could incorporate the perspectives of various stakeholders to provide a more comprehensive understanding of the in-service training needs for EFL teachers.

Including policy makers, program designers, school administrators, teacher trainers, Education faculty members, pre-service teachers, and language learners in research studies would offer valuable insights into the broader context of in-service training programs. Their perspectives can contribute to a more holistic understanding of the challenges, requirements, and effective strategies for enhancing language assessment practices in EFL education. Engaging these stakeholders in future research would help bridge the gap between theory and practice, leading to more informed and effective in-service training programs that address the needs and expectations of all parties involved in EFL teaching and learning.

Last but not least, future studies can combine quantitative and qualitative approaches with classroom observations to gain a more comprehensive understanding of EFL teachers' assessment practices. This would involve directly observing teachers' implementation of assessments, analyzing assessment samples, and gathering real-time data to complement the self-reported information obtained through questionnaires and interviews.

References

- Alderson, J. C. (2005). *Diagnosing foreign language proficiency. The interface between learning and assessment*. London, UK: Continuum.
- Al-Bahlani, S. M. (2019). *Assessment Literacy: A Study of EFL Teachers' Assessment Knowledge, Perspectives, and Classroom Behaviors*. Unpublished PhD Thesis, The University of Arizona, The United States.
- Alkharusi, H. (2011). Psychometric properties of the teacher assessment literacy questionnaire for preservice teachers in Oman. *Procedia-Social and Behavioral Sciences*, 29, 1614– 1624.
- Bachman, L. F. (2014). Ongoing challenges in language assessment. In A. J. Kunnan (Ed.), *the companion to language assessment* (1st ed., Vol. 3), (pp. 1586-1603). Oxford: John Wiley and Sons.
- Ballıdag, S. & Inan Karagül, B. (2021). Exploring the language assessment literacy of Turkish in-service EFL teachers. *Balıkesir Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*. <https://doi.org/10.31795/baunsobed.909953>
- Barduhn, S. & Johnson, J. (2005). Certification and professional qualifications. In A. Burns, & J. C. Richards (eds.) *Cambridge guide to second language teacher education*. (pp. 59-65). New York, NY: Cambridge University Press.
- Barbour, R. S., (2001). *Checklists for improving rigour in qualitative research: A case of the tail wagging the dog? BMJ*, 322(7294), 1115-1117.
- Bayrakçı, M. (2009). In-service teacher training in Japan and Turkey: A comparative Analysis of institutions and practices. *Australian Journal of Teacher Education*, 34(1), 10-22.
- Black, P. & Wiliam, D. (1998). Assessment and Classroom Learning, *Assessment In Education: Principles, Policy & Practice*, 5(1), 7-74.
- Boston, C. (2002). The concept of formative assessment. *Practical assessment, Research and Evaluation*, 8(9): 1–6. ERIC Clearinghouse on Assessment

and Evaluation.

- 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)
- Brown, J.D. (2001). *Using Surveys in Language Programs*. Cambridge: Cambridge University Press.
- Brown, H. D. (2003). *Language assessment: Principles and classroom practices*. Pearson Education.
- Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. New York: Pearson Education.
- Brown, H. D., & Abeywickrama, P. (2010). *Language assessment: Principles and Classroom Practices*. White Plains, NY: Pearson Education.
- Brown, J. D., & Bailey, K. M. (2008). Language testing courses: What are they in 2007. *Language Testing*, 25(3), 349–383.
- Burns, A., & Richards, J. C. (2009). Introduction. In A. Burns, & J. C. Richards (eds.) *Cambridge guide to second language teacher education*. (pp. 1-8). New York, NY: Cambridge University Press.
- Büyükkarlı, K. (2016). Identifying the areas for English language teacher development: A study of assessment literacy. *Pegem Eğitim ve Öğretim Dergisi*, 6(3), 333- 346.
- Campbell, J. L., Quincy, C., Osseman, J., & Pedersen, O. K. (2013). Coding in-depth semistructured interviews: Problems of unitization and intercoder reliability and agreement. *Sociological Methods & Research*, 42(3), 294-320.
- Carless, D. (2011). *From testing to productive student learning: Implementing formative assessment in Confucian-heritage settings*. London: Routledge
- Cirit, N. C. (2015). Assessing ELT Pre-Service Teachers via Web 2.0 Tools: Perceptions toward Traditional, Online and Alternative Assessment. *The*

- Turkish Online Journal of Educational Technology*, 14(3), 9-19.
- Clapham, C. (2000). Assessment And Testing. *Annual Review of Applied Linguistics*, 20,147. 161. doi:10.1017/s0267190500200093
- Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education (8th ed.)*. London: Routledge.
- Coombe, C. A. 1., Folse, K. S., & Hubley, N. J. (2007). *A practical guide to assessing English language learners*. Ann Arbor, Mich.: University of Michigan.
- Coombe, C. (2018). *An A to Z of Second Language Assessment: How Language Teachers Understand Assessment Concepts*. London, UK: British Council.
- Coskun, A., & Daloglu, A. (2010). Evaluating an English Language Teacher Education Program Through Peacock's Model. *Australian Journal of Teacher Education*, 35(6).
- Creswell, J. W. (2009). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (3rd ed.). Thousand Oaks, CA: Sage Publications.
- Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Boston, MA: Pearson.
- Crossman, J. (2007). The role of relationships and emotions in student perceptions of learning and assessment. *Higher Education Research and Development*, 26(3), 313-327.
- Csépes, I. (2013, November). *Language assessment literacy in the Hungarian context: What is it like and how is it developed in English teacher training?* Paper presented at the 4th CBLA SIG Symposium, Cyprus.
- Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives*, 8 (1).
- Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing*, 25(3), 327-347. <https://doi.org/10.1177/0265532208090156>.
- DeLuca, C., & Johnson, S. (2017). Developing assessment capable teachers in this

- age of accountability. *Assessment in Education: Principles, Policy & Practice*, 24 (2), 121-126, doi: 10.1080/0969594X.2017.1297010
- Deluca, C., & Klinger, D. A. (2010). Assessment literacy development: Identifying gaps in teacher candidates' learning. *Assessment in Education: Principles, Policy & Practice*, 17(4), 419-438. doi:10.1080/0969594x.2010.516643
- Dochy, F. (2009). The edumetric quality of new modes of assessment: Some issues and prospects. In G. Joughin (Ed.), *Assessment, Learning and Judgement in Higher Education* (pp. 85-114). New York: Springer.
- Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative and mixed methodologies*. Oxford: Oxford University Press.
- Duboc, A. P. M. (2009). Language assessment and the new literacy studies. *Lenguaje*, 37(1), 159-178.
- Fard, Z. R. and Tabatabaei, O. (2018). Investigating assessment literacy of EFL teachers in Iran. *Journal of Applied Linguistics and Language Research*, 5(3), 91-100.
- Field, A. (2018). *Discovering Statistics Using IBM SPSS Statistics* (5th ed.). London, UK: Sage Publications
- Fraenkel, J. R., & Wallen, N. E. (2005). *How to design and evaluate research in education*. (6th ed.). New York: Mc Graw Hill.
- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113 - 132.
- Giraldo, F. (2018). Language assessment literacy: implications for language teachers. *Profile: Issues in Teachers' Professional Development*, 20(1), 179-195. <https://doi.org/10.15446/profile.v20n1.62089>.
- Grossman, G. M., Onkol, P. E., & Sands, M. (2007). Curriculum reform in Turkish teacher education: Attitudes of teacher educators towards change in an EU candidate nation. *International Journal of Educational Development*, 27(2), 138-150.
- Guerin, E. M. C. (2010). *Initial findings from a pilot Italian study of foreign language*

- teachers' stated language assessment knowledge-base and needs*. Paper presented at the Lancaster University Postgraduate Conference in Linguistics & Language Teaching, Lancaster, the United Kingdom.
- Günel, M., & Tanriverdi, K., (2014). In-service teacher training from international and national perspectives: The retention and loss of institutional and academic memories. *Education & Science*, 39(175), 73-94.
- Han, T. & Kaya, H. İ. (2014). Turkish EFL teachers' assessment preferences and practices in the context of constructivist instruction. *Journal of Studies in Education*, 4(1), 77-93.
- Hasselgreen, A., Carlsen, C. & Helness, H. (2004). *European Survey of Language and Assessment Needs. Part One: General Findings*. Retrieved April 11, 2017 from <http://www.ealta.eu.org/documents/resources/survey-reportpt1.pdf>.
- Hatipoğlu, Ç. (2010). Summative Evolution 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.
- Haznedar, B. (2012). Perspectives on Pre-service Education of English Language Teachers in Turkish Primary School. In Y.Bayyurt & Y.Bektaş Çetinkaya (Eds.). *Research Perspectives on Teaching and Learning English in Turkey: Policies and Practices* (pp. 39-57). Peterlang.
- Herrera, L., & Macías, D. (2015). A call for language assessment literacy in the education and development of teachers of English as a foreign language. Colombia. *Appl. Linguist. J.*, 17(2), 302-312.
- Hudaya, D. W. (2017). Teachers' assessment literacy in applying principles of language assessment. *Proceedings of Education and Language International Conference*, 1(1), 247-260.

- Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*, 25, 328–402.
- Jeong, H. (2013). Defining assessment literacy: Is it different for language testers and non-language testers? *Language Testing*, 30(3), 345-362.
- Jin, Y. (2010). The place of language testing and assessment in the professional preparation of foreign language teachers in China. *Language Testing*, 27(4), 555–584.
- Johnson, R. B., & Christensen, L. (2004). *Educational research: Quantitative, qualitative, and mixed approaches*. Boston, MA: Pearson Education, Inc.
- Johnson, B., & Christensen, L. B. (2017). *Educational research: quantitative, qualitative, and mixed approaches* (6th ed.). Thousand Oaks, CA: SAGE Publications, Inc.
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33(7), 14-26.
- Kahraman-Özkurt, S. (2019). *Listening to English language teachers in Turkey: A Survey of their evaluations, preferences and needs regarding in-service training*. (Unpublished master's thesis). Gazi University, Ankara.
- Karagül, İ., B., Yüksel, D. & Altay, M. (2017). Assessment and grading practices of EFL teachers in Turkey. *International Journal of Language Academy*, 5(5) CUELT Special Issue, 168-174.
- Kaya, T., & Mede, E. (2021). Exploring language assessment literacy of EFL instructors in language preparatory programs. *Eğitim Fakültesi Dergisi*, 7(1), 163–189. https://doi.org/10.17932/iau.efd.2015.013/efd_v07i008
- Kiomrs, R., Abdolmehdi, R., & Naser, R. (2011). On the interaction of test washback and teacher assessment literacy: The case of Iranian EFL secondary school teachers. *English Language Testing*, 4(1), 156-161.
- Kırkgöz, Y. (2007). English language teaching in Turkey: Policy changes and their implementations. *RELC*, 38, 216-228.
- Koh, K. H. (2011). Improving teachers' assessment literacy through professional

- development. *Teaching Education*, 22(3). 255-276.
- 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).
- Küçüksüleymanoğlu, R. (2006). In-service training of ELT teachers in Turkey between 1998- 2005. *Eğitim Fakültesi Dergisi*, 9(2), 359-369.
- Lam, R. (2015). Language assessment training in Hong Kong: Implications for language assessment literacy. *Language Testing*. Published online. doi:10.1177/0265532214554321
- Lambert, D., & Lines, D. (2000). *Understanding Assessment: Purposes, Perception and Practice*. London: RoutledgeFalmer
- Lavrakas, P. J. (2008). *Encyclopedia of survey research methods* (Vols. 1-0). Thousand Oaks, CA: Sage Publications, Inc. DOI: 10.4135/9781412963947
- Leaph, K., Channy, M., & Chan, N. K. (2015). Cambodian ELT university practitioners' use of standardized tests for practice and assessment. *Language Education in Asia*, 6(1), 4-16.
- Lodico, M. G., Spaulding, D. T. & Voegtler K. H. (2006). *Methods in educational research: From theory to practice*. San Francisco, CA: Jossey-Bass Wiley.
- Lynch, B. K. (2003). *Language assessment and program evaluation*. Edinburgh Textbooks in Applied Linguistics Series (A. Davies & K. Mitchell, Eds.). Edinburgh: Edinburgh University Press.
- Mackey, A., & Gass, S. M. (2005). *Second language research: Methodology and design*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Maclellan, E. (2004). Initial knowledge states about assessment: Novice teachers' conceptualizations. *Teaching and Teacher Education*, 20(5), 523e535.
- Mahalingappa, L. J., & Polat, N. (2013). English language teacher education in Turkey: Policy vs academic standards. *European Journal of Higher Education*, 3(4), 371-383.

- Malone, M. E. (2013). The essentials of assessment literacy: Contrasts between testers and users. *Language Testing*, 30(3), 329–344.
- Mede, E., & Atay, D. (2017). English Language Teachers' Assessment Literacy: The Turkish Context. *Ankara Üniversitesi, TÖMER Dil Dergisi*, 168 (1), 43- 60.
- Mertler, C. A. (2004). Secondary teachers' assessment literacy: Does classroom experience make a difference? *American Secondary Education*, 33(1), 49–64.
- Mertler, C. A. (2009). Teachers' assessment knowledge and their perceptions of the impact of classroom assessment professional development. *Improving Schools*, 12(2), 101-113.
- Milli Eğitim Bakanlığı Öğretmen Yetiştirme ve Geliştirme Genel Müdürlüğü. (2023a). Hizmetiçi eğitim planları: 2001-2023 yıllarında planlanan faaliyet bilgileri. Retrieved January 20, 2023, from http://oygm.meb.gov.tr/www/icerik_goruntule.php?KNO=28
- Milli Eğitim Bakanlığı. (2023b). Öğretmenlerin hizmet içi eğitimleri "EBA" üzerinden yapılacak. Retrieved January 10, 2023, from: <http://www.meb.gov.tr/ogretmenlerin-hizmet-ici-egitimleri-eba-uzerinden-yapilacak/haber/20512/tr>
- Mirici, İ. H., & Pulatsü, G. (2022). Current trends in professional development: A quick guide for English language educators. *Journal for Educators, Teachers and Trainers*, 13(4). 317-328.
- MoNE. (1995). Hizmetiçi Eğitim Yönetmeliği [In-Service training regulations] Retrieved from <http://mevzuat.meb.gov.tr/html/51.html>
- MoNE. (2018). Professional development activities. Retrieved from <https://oygm.meb.gov.tr/www/mesleki-gelisim-tanitim/icerik/7>.
- Morrow, K. (Ed.), (2004). *Insights from the Common European Framework*. Cambridge, UK: Cambridge University Press.
- Muijs, D. (2004). *Doing quantitative research in education with SPSS*. SAGE

Publications Ltd.

Nagai, N. (2020). *CEFR-informed learning, teaching and assessment: A practical guide*. Singapore, Singapore: Springer.

Newsfield, T. (2006). Teacher development and assessment literacy. *Authentic Communication: Proceedings of the 5th Annual JALT Pan-SIG Conference*. (pp. 48- 73).

Ölmezer-Öztürk, E. & Aydın, B. (2018a). *Developing and Validating Language Assessment Knowledge Scale (LAKS) and Exploring the Assessment Knowledge of EFL Teachers*. Unpublished PhD Dissertation, Anadolu University, Turkey.

Ölmezer-Öztürk, E., & Aydın, B. (2018b). Investigating language assessment knowledge of EFL teachers. *Hacettepe University Journal of Education*. Advance online publication. doi: 10.16986/HUJE.2018043465

Özer, B. (2004). In-service training of teachers in Turkey at the beginning of the 2000s. *Journal of In-service education*, 30(1), 89-100.

Paterno, J. (2001). *Measuring success: A glossary of assessment terms*. In building cathedrals: Compassion for the 21st century.

Patton, M. Q. (2002). *Qualitative research and evaluation methods*. Thousand Oaks, CA: Sage.

Popham, W. J. (2004). Why assessment illiteracy is professional suicide. *Educational Leadership*, 62(1), 82-83.

Popham, W. J. (2006). Needed: A dose of assessment literacy. *Educational Leadership*, 63(6), 84–85.

Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory into Practice*, 48(1), 4–11

Purpura, J. E. (2016). Second and foreign language assessment. *Modern Language Journal*, 100 (Supplement 2016), 290-208. (Centennial Issue).

Richards, J. C. (1984). Language Curriculum Development. *RELC Journal*, 15(1), 1-29.

- Ryan, R. L. (1987). *The complete inservice staff development program*. USA:Prentice-Hall Inc.
- Salami, F. A., & Alharthi, R. M. (2022). Improving Language Assessment Literacy for In-Service Saudi EFL Teachers Arab World English Journal, 13 (3) 536-554. doi: <https://dx.doi.org/10.24093/awej/vol13no3.35>
- Saldaña, J. & Omasta, M. (2018). *Qualitative research: Analyzing life*. Thousand Oaks, CA: SAGE Publications.
- Scarino, A. (2013). Language assessment literacy as self-awareness: Understanding the role of interpretation in assessment and in teacher learning. *Language Testing*, 30(3), 309-327. doi: 10.1177/0265532213480128
- Siegel, M. A., & Wissehr, C. (2011). Preparing for the plunge: Preservice teachers' assessment literacy. *Journal of Science Teacher Education*, 22(4), 371-391.
- Stiggins, R. J. (1991). Relevant classroom assessment training for teachers. *Educational Measurement: Issues and Practice*, 10(1), 7-12.
- Stiggins, R.J. (1999). Evaluating classroom assessment training in teacher education programs. *Educational Measurement: Issues and Practice*, 18(1), 23–27.
- Stiggins, R. J. (2001). The unfulfilled promise of classroom assessment. *Educational Measurement: Issues and Practice*, 20(3), 5-15.
- Stiggins, R. J. (2002). Assessment crisis: The absence of assessment FOR learning. *Phi Delta Kappan*, 83(10), 758–766.
- Stiggins, R.J. (2008). *Assessment Manifesto: A Call for the Development of Balance Assessment Systems. A position paper published by the ETS Assessment Training Institute, Portland, Oregon.*
- Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Thousand Oaks, CA: Sage.
- Subedi, B. P. (2016). Using Likert Type Data in Social Science Research: Confusion, Issues and Challenges. *International Journal of Contemporary Applied Sciences*,

- 3(2), 36-49.
- Sultana, N. (2019). Language assessment literacy: An uncharted area for the English language teachers in Bangladesh. *Language Testing in Asia*, 9(1).
- Tamerer, R. B. (2019). *An investigation of Turkish pre-service EFL teachers' language assessment literacy* (Unpublished master's thesis). Kocaeli University.
- Taras, M., (2005), Assessment: Summative and Formative-Some Theoretical Reflections, *British Journal of educational Studies*, University of Sunderland, Blackwell Publishing Ltd, 53(4): pp. 466-478.
- Tashakkori, A., & Teddlie, C. (2003). *Handbook of Mixed Methods in Social & Behavioral Research*. Thousand Oaks: Sage
- Taylor, L. (2009). Developing assessment literacy. *Annual Review of Applied Linguistics*, 29, 21–36.
- Tesch, R. (1990). *Qualitative research: Analysis types and software tools*. Falmer, New York.
- Tierney, R. D. (2013). Fairness in classroom assessment. In J. H. McMillan (Ed.), *SAGE handbook of research on classroom assessment* (pp. 125e144). Thousand Oaks, CA: SAGE.
- 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), 18- 40.
- Uysal, H. H. (2012). Evaluation of an in-service training program for primary- school language teachers in Turkey. *Australian Journal of Teacher Education*, 37(7), 14-29.
- Uztosun, M.S. (2018). In-service teacher education in Turkey: English language teachers' perspectives. *Professional Development in Education*, 44(4), 557- 569.
- Vogt, K., & Tsagari, D. (2014). Assessment Literacy of Foreign Language Teachers: Findings of a European Study, *Language Assessment Quarterly*, 11(4), 374-402.
- Vogt, K., Tsagari, D., & Spanoudis, G. (2020): What Do Teachers Think They Want?

A Comparative Study of In-Service Language Teachers' Beliefs on LAL Training Needs, *Language Assessment Quarterly*, DOI: 10.1080/15434303.2020.1781128

- Vogt, W., P., Vogt, E., P., Gardner, D., C., & Lynne, M., H. (2014). *Selecting the Right Analyses for Your Data Quantitative, Qualitative, and Mixed Methods*. The USA, New York: The Guilford Press.
- Volante, L. & Fazio, X. (2007). Exploring teacher candidates' assessment literacy: Implications for teacher education reform and professional development. *Canadian Journal of Education*, 30(3), 749–770.
- Wach, A. (2012). Classroom-based language efficiency assessment: A challenge for EFL teachers. *Glottodidactica*, 39(1), 81-92.
- Webb, N. (2002, April). *Assessment literacy in a standards-based education setting*. A paper presented at the annual meeting of the American Educational Research Association, New Orleans, Louisiana.
- Wood, D. L. (2007). *Assessment for learning: Connecting students to their learning*. (Unpublished master's thesis).
- Xu, H. (2017). Exploring novice EFL teachers' classroom assessment literacy development: A three-year longitudinal study. *Asia-Pacific Educational Research*, 26(3-4), 219-226. doi:10.1007/s40299-017-0342-5
- Xu, Y., & Brown, G. T. L. (2016). Teacher assessment literacy in practice: A Reconceptualization. *Teaching and Teacher Education*, 58, 149–162. <https://doi.org/10.1016/j.tate.2016.05.010>
- 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), 133-158.
- Xu, Y., & Liu, Y. (2009). Teacher assessment knowledge and practice: A narrative inquiry of a Chinese college EFL teacher's experience. *TESOL Quarterly*, 43(3), 493-513.

- Yastıbaş, A., E. (2018). *Language assessment literacy of Turkish EFL instructors: A multiple case study*. Unpublished PhD Thesis, Atatürk University, Turkey.
- Yastıbaş, A. E., & Takkaç, M. (2018). Understanding language assessment literacy: Developing language assessment. *Journal of Language and Linguistic Studies*, 14(1), 178-193.
- Yüksek Öğretim Kurulu [Council of Higher Education (CoHE)] (2018). Yeni öğretmen yetiştirme lisans programları [New pre-service teacher training programs]. https://www.yok.gov.tr/Documents/Kurumsal/egitim_ogretim_dairesi/Yeni-Ogretmen-Yetistirme-Lisans-Programlari/Ingilizce_Ogretmenligi_Lisans_Programi.pdf.

APPENDIX-A: LAL Questionnaire

Dear Participant,

The following survey aims to find out training you received in language testing and assessment in teacher education programs and your training needs in this field. There are no 'right' or 'wrong' answers. Please be assured that all the information included in this survey is confidential.

Your answers will have a valuable contribution to the study.

Thank you very much for your participation.

Gamze Sarıyıldız Canlı

University of Health Sciences

Part I. General Information

1) Age:

2) Gender: Female Male

3) Education Level:

BA

MA

PhD

4) Undergraduate major:

ELT

English Linguistics

English Language and Literature

Linguistics

American Culture and Literature

Translation and Interpreting

Other: _____

5) Master's major (if applicable):

ELT

English Language and Literature

Curriculum and Instruction

Educational Administration

Other: _____

6) PhD major (if applicable):

ELT

English Language and Literature

Translation and Interpreting

Curriculum and Instruction

Other: _____

7) How many years have you taught English?

1-5 years

6-10 years

11 or more years

8) Type of school/institution you teach at

Public School

Private School

Private Language School

9) The current level of education that they are teaching

Primary School

Secondary School

High School

10) Which language is your first language?

11) Have you ever taken any Testing and Evaluation course before?

Yes

No

12) Have you ever taken any Testing and Evaluation course in ELT before?

Yes

No

Part II. Questions about the training in LTA

Please specify the amount of training you think you have received during your university education, and you need, respectively in the following four domains by ticking the box.

essays, presentations, homeworks)								
17) Using continuous type of assessment (e.g. quizzes)								
18) Using European Language Portfolio								
Knowledge of Testing and Assessment								
1) Informal/ Formal assessment								
2) Formative/ Summative assessment								
3) Norm /Criterion-referenced assessment								
4) Discrete point/Integrative testing								
5) Direct/Indirect testing								
6) Objective/Subjective testing								
7) Approaches to language testing (e.g. integrative, communicative, structuralist)								
8) Alternative assessment								
9) Computer-based testing								
Purposes of Testing								
1) Giving grades								
2) Finding out what needs to be learned/taught								
3) Placing students onto programs, courses, etc.								
4) Testing competence in a language								
5) Identifying what has been learned								
6) Measuring general ability to learn a foreign language								
7) Awarding final certificates								
Content and concepts of LTA								

1) Testing reading in English								
2) Different test items/task types to test reading in English								
3) Testing listening in English								
4) Different test items/task types to test listening in English								
5) Testing speaking in English								
6) Different test items/task types to test speaking in English								
7) Testing writing in English								
8) Different test items/task types to test writing in English								
9) Testing Grammar in English								
10) Different test items/task types to test grammar in English								
11) Testing Vocabulary in English								
12) Different test items/task types to test vocabulary in English								
13) Testing integrated language skills								
14) Testing pronunciation in English								
15) Different test items/question types to test pronunciation in English								
16) Practicality								
17) Reliability (e.g. rater, test administration, test, student-related)								
18) Validity (face, construct, criterion, content)								

19) Authenticity								
20) Washback								
21) Using statistics to study the quality of tests / assessment								
22) Alternatives in assessment (portfolios, conferences, interviews, observations, self/peer assessment)								

Appendix-B: Semi-Structured Interview Form in Turkish

- 1) Hangi bölümlerden mezun oldunuz (Lisans, Yüksek Lisans ve Doktora)?
- 2) Ne kadar zamandır öğretmenlik yapıyorsunuz?
- 3) Lisans eğitiminizde kaç istatistik, ölçme ve değerlendirme kursu aldınız?
- 4) Lisans eğitiminde aldığınız Ölçme ve Değerlendirme dersinin dil ölçme ve değerlendirmesinin gerekli bileşenlerini kapsadığını düşünüyor musunuz?
- 5) Bu derste ele alınan konulardan hangisinin bir sınıf öğretmeni olarak size EN ÇOK yardımcı olduğunu düşünüyorsunuz?
- 6) Bu dersi tekrar alsaydınız veya dil ölçme ve değerlendirmesinde ileri düzey bir ders alırsanız, hangi konuları öğrenmek isterdiniz?
- 7) Sınıflarınızda hangi değerlendirme yönteminin kullanılacağına nasıl karar veriyorsunuz?
- 8) Lisans eğitimindeki Dil Ölçme ve Değerlendirme eğitiminin Dil öğretmeni olarak sizin için yeterli olduğunu düşünüyor musunuz?
Evet
Hayır
Cevabınız Hayır ise, Dil Ölçme ve Değerlendirmenin hangi alanlarında eğitime ihtiyacınız olduğunu söyleyebilir misiniz?
- 9) Öğrenme ve öğretme kalitesini artırmak için dil öğretmenlerinin Dil Ölçme ve Değerlendirme alanında yetkin olmasının önemli olduğuna inanıyor musunuz?
- 10) Bir dil öğretmeni olarak sınav ve değerlendirme uygulamalarıyla ilgili ne tür sorunlar ve zorluklar yaşadınız (ör. Testlerin hazırlanması ve uygulanması, öğrenciler, güvenilirlik)? Bu sorunları nasıl aşarsınız?
- 11) Konuyla ilgili sorularınız veya yorumlarınız var mı?

Appendix-C: Semi-Structured Interview Form in English

- 1) Which departments did you graduate from (BA., MA and PhD)?
- 2) How long have you been teaching?
- 3) How many statistics and language testing and assessment course have you taken in your undergraduate training?
- 4) Do you think that Testing and Evaluation Course that you took in undergraduate training covers necessary components of language testing and assessment?
- 5) Of the topics covered in this course, which topic/s do you think is/are the MOST helpful to you as a classroom teacher?
- 6) If you were to take this course again, or take an advanced course in language testing and assessment, what topics would you like to learn about?
- 7) How do you decide which assessment method to use in your classrooms?
- 8) Do you think that LTA training in undergraduate education is adequate for you as a language teacher?

Yes

No

If the answer is No, could you state in which areas of LTA do you think you need training?

- 9) Do you believe it is important for language teachers to be competent at LTA in order to increase the quality of learning and teaching?
- 10) What kind of problems and difficulties you have experienced related to testing and assessment practices as a language teacher (e.g. preparation and administration of tests, students, reliability)? How do you overcome these problems?
- 11) Do you have any questions or comments about the subject?

Appendix-D: Interview Schedule

	Date	Time	Pseudonym
1.	10.10.21	16:00	Seda
2..	11.10.21	11:30	Ahmet
3.	11.10.21	14:30	Alican
4.	13.10.21	10:00	Seher
5.	13.10.21	18:30	Alper
6.	14.10.21	12:00	Melis
7.	16.10.21	19:00	Selin
8.	17.10.21	17:00	Berkin
9.	18.10.21	11:00	Doğa
10.	20.10.21	09:30	Fatma
11.	20.10.21	10:00	Dilan
12.	20.10.21	11:30	Merve
13.	20.10.21	13:00	Cenk
14.	21.10.21	16:00	Ada
15.	21.10.21	18:00	Gizay
16.	22.10.21	09:15	Mustafa
17.	22.10.21	09:45	Ayla
18.	22.10.21	19:45	Okan
19.	23.10.21	08:45	Sezen
20.	23.10.21	15:30	Eda

Appendix-E: List of Codes for the Interview Form

On the left column, themes are provided and on the right column, the codes and sub-codes are presented. The frequency of the codes is provided in parenthesis and sub-codes which are less frequent than 3 are not given.

Education_level	<ul style="list-style-type: none"> ▪ BA [14] <ul style="list-style-type: none"> ❖ ELT [10] ❖ ELIT [4] ▪ MA [6] <ul style="list-style-type: none"> ❖ ELT [3] ❖ Curriculum and Instruction [3]
Years_of_Experience	<ul style="list-style-type: none"> ▪ 1-5 years [6] ▪ 6-10 years [5] ▪ 11 or more years [9]
The_Number_of_Statistics_and_LTA_Course	<ul style="list-style-type: none"> ▪ 1 LTA course [20] ▪ 1 Statistics course [3]
Sufficiency_of_Necessary_Components	<ul style="list-style-type: none"> ▪ Sufficient [5] ▪ Insufficient [15]
Helpful_Topics	<ul style="list-style-type: none"> ▪ Adapting prepared tests [4] ▪ Alternative assessment [5] ▪ Objective testing [3] ▪ Formative/Summative assessment [3] ▪ Testing reading [3] ▪ Practicality [4] ▪ Validity [5] ▪ Informal/Formal assessment [4]
Topics_to_Learn_More_about	<ul style="list-style-type: none"> ▪ Statistics [6] ▪ Giving feedback [4] ▪ Reliability [3] ▪ Alternative assessment [5] ▪ Computer based testing [4] ▪ Testing Integrated Language Skills [3] ▪ Validity [3] ▪ Approaches to Language Testing [3]
Deciding_which_method_to_use	<ul style="list-style-type: none"> ▪ Talking to colleagues [4] ▪ Doing research on the internet [5] ▪ Reading assessment books [7] ▪ Remembering previous knowledge and experience [5]
Sufficiency_of_LTA_training_in_Undergraduate_Education	<ul style="list-style-type: none"> ▪ Yes [6] ▪ No [14]

	<ul style="list-style-type: none"> ❖ Testing speaking [3] ❖ Testing listening [2] ❖ Testing vocabulary [3] ❖ Testing grammar [4] ❖ Statistics [5] ❖ Using ELP [4]
The_Importance_of_LTA_competence	<ul style="list-style-type: none"> ▪ Highly crucial [13] <ul style="list-style-type: none"> ❖ Relation among testing, learning, and teaching [7] ❖ Using appropriate assessment methods [5] ❖ Giving reliable information about students [4] ▪ To some extent [4] <ul style="list-style-type: none"> ❖ Informing students [3] ▪ Not crucial [3] <ul style="list-style-type: none"> ❖ No direct relation among the quality of testing, teaching and learning [3]
Problems_and_Difficulties_in_LTA	<ul style="list-style-type: none"> ▪ Preparation of tests [10] ▪ Administration of tests [5] ▪ Learners related [4] ▪ Reliability [3] ▪ Giving feedback [4] ▪ Validity [3]
Ways_of_Handling_Problems	<ul style="list-style-type: none"> ▪ Consulting to another colleague [9] ▪ Conducting research and reading related things on the internet [8] ▪ Ignoring problems [3]
Further_Comments_and_Questions	<ul style="list-style-type: none"> ▪ Ways to develop language assessment knowledge [3]

APPENDIX-F: Ethics Committee Approval



T.C.
HACETTEPE ÜNİVERSİTESİ REKTÖRLÜĞÜ
Rektörlük

Tarih: 01/07/2021 08:21
Ser. E-35853172-101.02.02-00001637630



00001637630

Sayı : E-35853172-101.02.02-00001637630
Konu : Gamze SARIYILDIZ CANLI (Etik Komisyon İzni)

1.07.2021

EĞİTİM BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜNE

İlgi: 26.04.2021 tarihli ve E-51944218-101.02.02-00001552424 sayılı yazı.

Enstitünüz Yabancı Diller Eğitimi Anabilim Dalı İngiliz Dili Eğitimi Doktora programı öğrencisi **Gamze SARIYILDIZ CANLI**'nın **Dr. Öğr. Üyesi İsmail Fırat ALTAY** danışmanlığında yürüttüğü "**Türkiye'deki Hizmetçi İngilizce Öğretmenlerinin Dil Değerlendirme Okuryazarlık Düzeyi ve İhtiyaçları/Language Assesment Literacy Level and Needs Of In-Service EFL Teachers in Turkey**" başlıklı tez çalışması Üniversitemiz Senatosu Etik Komisyonunun **22 Haziran 2021** tarihinde yapmış olduğu toplantıda incelenmiş olup, etik açıdan uygun bulunmuştur.

Bilgilerinizi ve gereğini saygılarımla rica ederim.

e-İmzalıdır
Prof. Dr. Vural GÖKMEN
Rektör Yardımcısı

Bu belge güvenli elektronik imza ile imzalanmıştır.

Belge Doğrulama Kodu: 203C16E4-B035-41C4-88F4-77AFA90FC760

Belge Doğrulama Adresi: <https://www.turkiye.gov.tr/hu-cbys>

Adres: Hacettepe Üniversitesi Rektörlük 06100 Sıhhiye-Ankara

Bilgi için: Sevdâ TOPAL

E-posta: yuzimd@hacettepe.edu.tr İnternet Adresi: www.hacettepe.edu.tr Elektronik

Bilgisayar İşletmeni

Aç: www.hacettepe.edu.tr

Telefon: 0 (312) 305 3001-3002 Faks: 0 (312) 311 9992

Telefon: 03123051008

Kep: hacettepeuniversitesi@hs01.kep.tr



APPENDIX-G: Declaration of Ethical Conduct

I hereby declare that...

- I have prepared this thesis in accordance with the thesis writing guidelines of the Graduate School of Educational Sciences of Hacettepe University;
- all information and documents in the thesis/dissertation have been obtained in accordance with academic regulations;
- all audio visual and written information and results have been presented in compliance with scientific and ethical standards;
- in case of using other people's work, related studies have been cited in accordance with scientific and ethical standards;
- all cited studies have been fully and decently referenced and included in the list of References;
- I did not do any distortion and/or manipulation on the data set,
- and **NO** part of this work was presented as a part of any other thesis study at this or any other university.

(06)/(07)/(2023)

(Signature)

Gamze SARIYILDIZ CANLI

APPENDIX-H: Dissertation Originality Report

06/07/2023
 HACETTEPE UNIVERSITY
 Graduate School of Educational Sciences
 To The Department of Foreign Language Education

Thesis Title: LANGUAGE ASSESSMENT LITERACY LEVEL AND NEEDS OF IN-SERVICE EFL TEACHERS IN TURKEY

The whole thesis that includes the *title page, introduction, main chapters, conclusions and bibliography section* is checked by using **Turnitin** plagiarism detection software take into the consideration requested filtering options. According to the originality report obtained data are as below.

Time Submitted	Page Count	Character Count	Date of Thesis Defense	Similarity Index	Submission ID
06/07/2023	233	369,121	06/07/2023	11%	2081079503

Filtering options applied:

1. Bibliography excluded
2. Quotes included
3. Match size up to 5 words excluded

I declare that I have carefully read Hacettepe University Graduate School of Educational Sciences Guidelines for Obtaining and Using Thesis Originality Reports; that according to the maximum similarity index values specified in the Guidelines, my thesis does not include any form of plagiarism; that in any future detection of possible infringement of the regulations I accept all legal responsibility; and that all the information I have provided is correct to the best of my knowledge.

I respectfully submit this for approval.

Name Lastname: Gamze Sarıyıldız Canlı
Student No.: N18144432
Department: Foreign Language Education
Program: English Language Teaching
Status: Masters Ph.D. Integrated Ph.D.

Signature

ADVISOR APPROVAL

APPROVED
 (Assist. Prof. Dr. İsmail Firat ALTAY)

APPENDIX-I: Yayınlama ve Fikrî Mülkiyet Hakları Beyanı

Enstitü tarafından onaylanan lisansüstü tezimin/raporumun tamamını veya herhangi bir kısmını, basılı (kâğıt) ve elektronik formatta arşivleme ve aşağıda verilen koşullarla kullanıma açma iznini Hacettepe Üniversitesine verdiğimi bildiririm. Bu izinle Üniversiteye verilen kullanım hakları dışındaki tüm fikri mülkiyet haklarım bende kalacak, tezimin tamamının ya da bir bölümünün gelecekteki çalışmalarda (makale, kitap, lisans ve patent vb.) kullanım hakları bana ait olacaktır.

Tezin kendi orijinal çalışmam olduğunu, başkalarının haklarını ihlal etmediğimi ve tezimin tek yetkili sahibi olduğumu beyan ve taahhüt ederim. Tezimde yer alan telif hakkı bulunan ve sahiplerinden yazılı izin alınarak kullanılması zorunlu metinlerin yazılı izin alınarak kullandığımı ve istenildiğinde suretlerini Üniversiteye teslim etmeyi taahhüt ederim.

Yükseköğretim Kurulu tarafından yayınlanan "**Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge**" kapsamında tezim aşağıda belirtilen koşullar haricince YÖK Ulusal Tez Merkezi / H.Ü. Kütüphaneleri Açık Erişim Sisteminde erişime açılır.

- Enstitü/Fakülte yönetim kurulu kararı ile tezimin erişime açılması mezuniyet tarihinden itibaren 2 yıl ertelenmiştir. ⁽¹⁾
- Enstitü/Fakülte yönetim kurulunun gerekçeli kararı ile tezimin erişime açılması mezuniyet tarihinden itibaren ... ay ertelenmiştir. ⁽²⁾
- Tezimle ilgili gizlilik kararı verilmiştir. ⁽³⁾

06 /07/2023

(imza)

Gamze SARIYILDIZ CANLI

"Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge"

- (1) Madde 6.1. Lisansüstü teze ilgili patent başvurusu yapılması veya patent alma sürecinin devam etmesi durumunda, tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulu iki yıl süre ile tezinerişime açılmasının ertelenmesine karar verebilir.
- (2) Madde 6.2. Yeni teknik, materyal ve metotların kullanıldığı, henüz makaleye dönüşmemiş veya patent gibi yöntemlerle korunmamış ve internette paylaşılması durumunda 3. şahıslara veya kurumlara haksız kazanç; imkânı oluşturabilecek bilgi ve bulguları içeren tezler hakkında tez danışmanın önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulunun gerekçeli kararı ile altı ay aşmamak üzere tezin erişime açılması engellenebilir.
- (3) Madde 7.1. Ulusal çıkarları veya güvenliği ilgilendiren, emniyet, istihbarat, savunma ve güvenlik, sağlık vb. konulara ilişkin lisansüstü tezlerle ilgili gizlilik kararı, tezin yapıldığı kurum tarafından verilir*. Kurum ve kuruluşlarla yapılan işbirliği protokolü çerçevesinde hazırlanan lisansüstü tezlere ilişkin gizlilik kararı ise, ilgili kurum ve kuruluşun önerisi ile enstitü veya fakültenin uygun görüşü üzerine üniversite yönetim kurulu tarafından verilir. Gizlilik kararı verilen tezler Yükseköğretim Kuruluna bildirilir.
Madde 7.2. Gizlilik kararı verilen tezler gizlilik süresince enstitü veya fakülte tarafından gizlilik kuralları çerçevesinde muhafaza edilir, gizlilik kararının kaldırılması halinde Tez Otomasyon Sistemine yüklenir
*Tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulu tarafından karar verilir

