



REPUBLIC OF TÜRKİYE
ÇANAKKALE ONSEKİZ MART UNIVERSITY
SCHOOL OF GRADUATE STUDIES

DEPARTMENT OF FOREIGN LANGUAGES EDUCATION
ENGLISH LANGUAGE TEACHING PROGRAM

**ENGLISH LANGUAGE TEACHERS' ASSESSMENT LITERACY:
IDENTIFYING ASSESSMENT NEEDS AND BELIEFS OF EFL
TEACHERS AND TEACHER CANDIDATES**

PHD THESIS

PINAR ÇANKAYA

Supervisor
PROF. DR. DİNÇAY KÖKSAL

ÇANAKKALE – 2022



REPUBLIC OF TÜRKİYE
ÇANAKKALE ONSEKİZ MART UNIVERSITY
SCHOOL OF GRADUATE STUDIES

DEPARTMENT OF FOREIGN LANGUAGES EDUCATION
ENGLISH LANGUAGE TEACHING PROGRAM

**ENGLISH LANGUAGE TEACHERS' ASSESSMENT LITERACY:
IDENTIFYING ASSESSMENT NEEDS AND BELIEFS OF EFL TEACHERS
AND TEACHER CANDIDATES**

PHD THESIS

PINAR ÇANKAYA

Supervisor
PROF. DR. DİNÇAY KÖKSAL

ÇANAKKALE – 2022



T.C.
ÇANAKKALE ONSEKİZ MART ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ



Pınar ÇANKAYA tarafından Prof. Dr. Dinçay KÖKSAL yönetiminde hazırlanan ve **25/08/2022** tarihinde aşağıdaki jüri karşısında sunulan “**English Language Teachers’ Assessment Literacy: Identifying Assessment Needs and Beliefs of EFL Teachers and Teacher Candidates**” başlıklı çalışma, Çanakkale Onsekiz Mart Üniversitesi Lisansüstü Eğitim Enstitüsü **Yabancı Diller Eğitimi Anabilim Dalı**’nda **DOKTORA TEZİ** olarak oy birliği ile kabul edilmiştir.

Jüri Üyeleri

İmza

Prof. Dr. Dinçay KÖKSAL

(Danışman)

Prof. Dr. Aysun YAVUZ

Prof. Dr. Bülent GÜVEN

Doç. Dr. Ömer Gökhan ULUM

Dr. Öğretim Üyesi Eda DURUK

Tez No : 10321338

Tez Savunma Tarihi : 25/08/2022

.....
Doç. Dr. Yener PAZARCIK

Enstitü Müdürü

.././2022

ETİK BEYAN

Çanakkale Onsekiz Mart Üniversitesi Lisansüstü Eğitim Enstitüsü Tez Yazım Kuralları'na uygun olarak hazırladığım bu tez çalışmasında; tez içinde sunduğum verileri, bilgileri ve dokümanları akademik ve etik kurallar çerçevesinde elde ettiğimi, tüm bilgi, belge, değerlendirme ve sonuçları bilimsel etik ve ahlak kurallarına uygun olarak sunduğumu, tez çalışmasında yararlandığım eserlerin tümüne uygun atıfta bulunarak kaynak gösterdiğimi, kullanılan verilerde herhangi bir değişiklik yapmadığımı, bu tezde sunduğum çalışmanın özgün olduğunu, bildirir, aksi bir durumda aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi taahhüt ve beyan ederim.

In this thesis that I prepared in accordance with Thesis Guidelines of Çanakkale Onsekiz Mart University School of Graduate Studies; I hereby declare that I have obtained the data, information and documents I presented in the thesis within the framework of academic and ethical rules, I have presented all information, documents, evaluations and results in accordance with scientific ethics and ethical rules, that I refer to all the works I have used in my thesis, and that I have not made any changes to the data used and that the work I have presented in this thesis is original, I undertake and declare that I accept all the loss of rights that may arise against me in the event of a contrary.

Pınar ÇANKAYA

25/08/2022

ACKNOWLEDGEMENTS

I am deeply grateful to the many people for their invaluable assistance during my PhD journey sharing their critics, opinions and knowledge to enhance my study. Firstly, I would like to thank to my supervisor Prof. Dr. Dinçay KÖKSAL for his support and encouragement without which this dissertation would never go forward. Secondly, I would like to express my deepest gratitude to Prof. Dr. Aysun Yavuz and Prof. Dr. Bülent Güven for their scientific guidance, insights and recommendations during my thesis meetings. I also feel grateful to my defence jury members Assoc. Prof. Dr. Ömer Gökhan Ulum and Asst. Prof. Dr. Eda Duruk for their critical comments and constructive feedback providing a wealth of insights for improving the study. Besides, I am particularly indebted to Prof. Dr. Sevim İnal and the participants in my study for their invaluable assistance and cooperation without which this study would never have been completed. Further, I owe much to my colleagues for their careful reading, valuable suggestions and contributions during the development of data collection instruments. Also, I feel the necessity to express my thanks to my Ethic Committee for giving permission to conduct the study. Lastly, my sincere and special thanks go to my little “*Honey*” – Yağmur Asel- whose love and smile provided me with endless encouragement and force. She deserves my deepest thanks and kisses.

Pınar ÇANKAYA

Çanakkale, 2022

ÖZET

İNGİLİZCE ÖĞRETMENLERİNİN ÖLÇME OKURYAZARLIĞI: YABANCI DİL İNGİLİZCE ÖĞRETMENLERİNİN VE ÖĞRETMEN ADAYLARININ ÖLÇME İHTİYAÇLARI VE İNANIŞLARINI BELİRLEME

Pınar ÇANKAYA

Çanakkale Onsekiz Mart Üniversitesi

Lisansüstü Eğitim Enstitüsü

Yabancı Diller Eğitimi Anabilim Dalı Doktora Tezi

Danışman: Prof. Dr. Dinçay KÖKSAL

25/08/2022, 190

Çeşitli paydaşlar arasında dil değerlendirme okuryazarlığına artan ilgi, öğrencilerin öğrenmelerini de geliştirdiğine yaygın olarak inanılan etkin ve verimli bir değerlendirme için gerekli olan temel bilgi, beceri ve yeterliliklere dikkat çekmiştir. Bu kapsamda, bu çalışma hem nitel hem de nicel verilerin toplandığı bir karma yöntem araştırma tasarımı aracılığıyla İngilizce öğretmenlerinin ve öğretmen adaylarının temel ölçme değerlendirme ihtiyaçlarını, düzeylerini ve inanışlarını belirlemeye çalışmaktadır. Bu nedenle, Fulcher (2012) tarafından geliştirilen anket uyarlanarak yarı yapılandırılmış, odak grupları ve takip görüşmelerini de içeren bir çalışma yapabilmek adına konuyu araştırmak için 8 araştırma sorusu oluşturulmuştur.

Sonuçlara bakıldığında, nitel ve nicel yolla elde edilen veriler; her iki katılımcı grubunun da ölçme değerlendirme seviyesinin iyi düzeyde olduğunu, bununla birlikte ölçme değerlendirme eğitim ihtiyaçlarının da orta düzeyde olduğunu göstermiştir. Özellikle İngilizce öğretmenlerinin “üretme becerilerini (konuşma/yazma) değerlendirme, sonuçları yorumlama, geçerlilik ve güvenilirliği sağlama” konularında eğitime yüksek ihtiyaç duyduğu tespit edilmişken, öğretmen adayları ise “istatistik kullanımı, test özelliklerini ve şablonunu yazma, dilde ölçme ve değerlendirmenin tarihsel gelişimi, geçerlilik ve güvenilirliği sağlama” alanlarında eğitime ihtiyaçları olduğunu belirtmişlerdir.

Nitel verilerle dilde ölçme ve değerlendirme inanışları sorgulanırken, bulgular İngilizce öğretmenleri katılımcılarının öğrenme odaklı değerlendirme ve alternatif değerlendirme yöntemlerini savunmalarına rağmen, uygulamada hem teorik hem de pratik anlamda yeterli düzeyde olmadıklarını ve daha fazla eğitime ihtiyaç duyduklarını göstermiştir. Son olarak, alana katkı sağlamak adına, yabancı dilde ölçme ve değerlendirme okuryazarlığı gelişimi için bir eğitim programı grafiği tasarlanmış ve araştırmanın ana sonuçlarının bir sentezi, hem pratik hem de teorik yönlerden çıkarımlar ve önerilerle birlikte çalışma sonlanmıştır.

Anahtar Kelimeler: Yabancı Dilde Ölçme ve Değerlendirme, Ölçme ve Değerlendirme Okuryazarlığı, Ölçme ve Değerlendirme İnançları

ABSTRACT
ENGLISH LANGUAGE TEACHERS' ASSESSMENT LITERACY: IDENTIFYING
ASSESSMENT NEEDS AND BELIEFS OF EFL TEACHERS AND TEACHER
CANDIDATES

Pınar ÇANKAYA

Çanakkale Onsekiz Mart University

School of Graduate Studies

Doctoral Dissertation in English Language Teaching

Supervisor: Prof. Dr. Dinçay KÖKSAL

25/08/2022, 190

A growing interest in Language Assessment Literacy (LAL, henceforth) across various stakeholders has drawn attention to the basic knowledge, skills and competences necessary for a sound and efficient assessment which is widely believed to enhance student learning. Within this scope, this study attempts to identify the main LAL needs, levels and beliefs of in-service EFL teachers and pre-service EFL teacher candidates through a mixed methods research design where both qualitative and quantitative data have been collected. To this end, 8 research questions were posed to investigate the issue by conducting a survey study including a questionnaire developed by Fulcher (2012) and semi structured, focus groups, and follow up interviews.

Doing so the current study attempts to ascertain how EFL teachers and EFL teacher candidates perceive their LAL levels and needs along with their underlying beliefs about language assessment. As displayed in the findings, the study unveiled that both groups of participants were found to have good level of LAL and moderate need for further training. More specifically, in-service EFL teachers were found to be in high need for “*rating productive skills (speaking/writing), interpreting scores, validation and reliability* based on the self responses obtained through the survey and interviews while pre-service EFL teachers noted their high need for further training in the following items: using statistics, writing test specifications/blueprints, history of language testing, validation and reliability.

When scrutinized closely, certain eye catching results are worth of mentioning. Surprisingly, as qualitative data suggested, even though in-service and pre-service EFL teachers support ‘learning oriented assessment’ and ‘alternative assessment’, they self reported that they need further training for alternative assessment techniques to put their theoretical knowledge into practice. Notably, in order to contribute to the field, a framework for improving language assessment literacy has been developed having identified the basic philosophical, practical and developmental concerns. Finally, a synthesis of the main conclusions of the research is provided in the final chapter along with implications and suggestions from both practical and theoretical aspects.

Keywords: Language Testing and Assessment, Assessment Literacy, Assessment Beliefs

TABLE OF CONTENTS

	Page No
JÜRİ ONAY SAYFASI.....	i
ETİK BEYAN.....	ii
ACKNOWLEDGEMENTS.....	iii
ÖZET	iv
ABSTRACT	vi
TABLE OF CONTENTS.....	viii
SYMBOLS AND ABBREVIATIONS	xii
LIST OF TABLES	xiii
LIST OF FIGURES	xv
CHAPTER ONE	
INTRODUCTION	
1	
1.1. Introduction.....	1
1.2. Statement of the Problem.....	2
1.3. Background of the Study.....	3
1.4. Purpose of the Study.....	5
1.5. Research Questions.....	5
1.6. Significance of the Study.....	6
1.7. Context of the Study.....	8
1.8. Limitations of the Study.....	9
1.9. Definition of the Terms.....	10
1.10 Organization of Thesis.....	11
1.11 Chapter Summary.....	11
CHAPTER TWO	
LITERATURE REVIEW	
13	
2.1. Chapter Introduction.....	13
2.2. Theoretical Framework.....	13

2.1.1. Rationale of Research.....	13
2.3. Historical Background.....	16
2.3.1. Traditional Language Assessment versus Social Constructivist.....	19
2.4. Teacher Beliefs.....	20
2.5. Language Assessment Literacy.....	26
2.5.1. Defining Qualities of Language Assessment Literacy.....	29
2.5.2. Relevant Research Studies.....	33
2.6. Assessment Literacy Training in Pre service Teacher Education Programs.....	37
2.6.1. Assessment Courses during Pre-Service Language Teacher Education in Türkiye.....	38
2.6.2. Language Assessment Practices in Turkish Education System.....	40
2.6.3. New Trends and Recent Language Assessment Challenges during Covid-19.....	41
2.7. Chapter Summary.....	45
CHAPTER THREE	
METHODOLOGY	
3.1. Chapter Introduction.....	47
3.2. The Purpose of the Study.....	47
3.3. Research Questions.....	48
3.4. Research Design and Rationale.....	49
3.5. Research Site and Participants.....	49
3.6. Data Collection Procedure.....	53
3.7. Development of the Instruments and Piloting Process.....	54
3.7.1. Development of the Questionnaire.....	54
3.7.2. Development of the Interview Questions.....	55
3.7.3. Course Documents.....	57
3.8. Data Analysis.....	57
3.8.1. Qualitative Data.....	57
3.8.2. Quantitative Data.....	59
3.9. Validity and Reliability Concerns.....	59
3.10 Ethical Considerations.....	61
3.11 Chapter Summary.....	62

CHAPTER FOUR
RESULTS AND DISCUSSION 63

4.1. Chapter Introduction.....	63
4.2. Results and Discussion for Research Question 1.....	63
4.3. Results and Discussion for Research Question 2.....	71
4.4. Results and Discussion for Research Question 3.....	80
4.4.1. Results and Discussion for Assessment Beliefs.....	82
4.4.2. Results and Discussion for Interview Question 1 and Question 9.....	82
4.4.3. Results and Discussion for Interview Question 2, Question 3, Question 5.....	89
4.4.4. Results and Discussion for Interview Question 6 and Question 7.....	98
4.4.5. Results and Discussion for Interview Question 4 and Question 8.....	103
4.5. Results and Discussion for Research Question 4.....	106
4.6. Results and Discussion for Research Questions 5 and 6.....	115
4.7. Results and Discussion for Research Questions 7.....	122
4.8. Results and Discussion for Research Questions 8.....	129
4.9. Chapter Summary.....	137

CHAPTER FIVE
CONCLUDING REMARKS AND IMPLICATIONS 138

5.1. Chapter Introduction.....	138
5.2. Summary of the Study.....	138
5.2.1. Purpose of the Study.....	138
5.2.2. Summary of the Methodology.....	139
5.2.3. Summary of the Results.....	140
5.3. Conclusions.....	146
5.3.1. Conclusions Associated with Pre-service Training.....	146
5.3.2. Conclusions Associated with Language Assessment Literacy Development.....	154
5.3.3. Conclusions Associated with Language Assessment Beliefs.....	158
5.4. Implications.....	164
5.4.1. Practical and Pedagogical Implications.....	165
5.4.2. Theoretical implications.....	167

5.4.3. Methodological Implications.....	167
5.5. Limitations of the Study.....	168
5.5.1. Conceptual Limitations	168
5.5.2. Methodological Limitations	169
5.5.3. Research Design Limitations	170
5.6. New Research Avenues for Future Studies.....	171
REFERENCES.....	173
APPENDICES.....	I
APPENDIX 1. CONSENT FORM.....	I
APPENDIX 2. OPEN ENDED INTERVIEW QUESTIONS 1 st VERSION.....	II
APPENDIX 3. OPEN ENDED INTERVIEW QUESTIONS 2 nd VERSION.....	IV
APPENDIX 4. LANGUAGE ASSESSMENT LITERACY INVENTORY.....	VI
APPENDIX 5. A SAMPLE GUIDELINE FOR TEACHING LAL.....	IX
APPENDIX 6. ETİK KURUL ONAYI.....	X
APPENDIX 7. PERMISSION FOR DATA TOOLS.....	XI
APPENDIX 8. PERMISSION FOR CONDUCTING RESEARCH.....	XII
CURRICULUM VITAE.....	XIX

SYMBOLS AND ABBREVIATIONS

CEFR	Common European Framework of Reference for Languages
CoHE	Council of Higher Education
ÇOMÜ	Çanakkale Onsekiz Mart Üniversitesi
EFL	English as a Foreign Language
ELT	English Language Teaching
ELTE	English Language Testing and Evaluation
ICT	Information and Communication Technology
IQ	Interview Question
LAL	Language Assessment Literacy
LTA	Language Testing and Assessment
m	Mean
MoNE	Ministry of National Education
n	Number
RQ	Research Question
SD	Standard deviation
SPSS	Statistical Package for Social Sciences
YDS	Yabancı Dil Sınavı (Foreign Language Examination)
YKS	Yükseköğretim Kurumları Sınavı
%	percent

LIST OF TABLES

Table No	Name of Table	Page No
Table 1	The information of the pre-service EFL teachers	50
Table 2	Background information about in-service EFL teachers	51
Table 3	Interviewed in-service and pre-service EFL teachers' profile	52
Table 4	Pre-service EFL teachers' perceived LAL levels	65
Table 5	Pre-service EFL teachers' perceived LAL needs	68
Table 6	In-service EFL teachers' perceived LAL needs	72
Table 7	In-service EFL teachers' perceived LAL levels	77
Table 8	A comparison of LAL needs and levels of all participants	79
Table 9	Main findings identified based on pre-service EFL teachers' responses in IQ1	82
Table 10	Main findings identified based on in-service EFL teachers' responses in IQ1	84
Table 11	Perspectives of participants for students responsibility in language assessment	87
Table 12	Main themes identified in IQ2 based on pre-service EFL teachers responses	90
Table 13	Main themes identified in IQ2 based on in-service EFL teachers responses	91
Table 14	Main themes identified in IQ3 based on in-service EFL teachers responses	93
Table 15	Main themes identified in IQ3 based on pre-service EFL teachers responses	94
Table 16	Main themes identified in IQ5 based on in-service EFL teachers responses	96
Table 17	Main themes identified in IQ3 based on pre-service EFL teachers responses	97
Table 18	Main themes identified in IQ6 based on in-service EFL teachers responses	98

Table 19	Main themes identified in IQ6 based on pre-service EFL teachers responses	99
Table 20	Main themes identified in IQ7 based on in-service EFL teachers responses	101
Table 21	Main themes identified in IQ7 based on pre-service EFL teachers responses	103
Table 22	Main themes identified in IQ4 based on all participants' responses	104
Table 23	A summary of training, LAL level, needs and language assessment course views of participants	107
Table 24	A comparison of training and perceived LAL level of participants	113
Table 25	A comparison of training and perceived LAL need of participants	114
Table 26	Pre-service EFL teachers' perceived importance levels of LAL topics	116
Table 27	In-service EFL teachers' perceived importance levels of LAL topics	117
Table 28	The perceptions of pre-service EFL teachers about LAL topics	118
Table 29	The perceptions of in-service EFL teachers about LAL topics	120
Table 30	The main findings for IQ10	122
Table 31	Main themes identified for IQ4 and IQ5 in "needs for assessment" category	124
Table 32	Main themes identified for "online language assessment" by pre-service EFL teachers	130
Table 33	Main themes identified for "online language assessment" by in-service EFL teachers	133

LIST OF FIGURES

Figure No	Name of Figure	Page No
Figure 1	The test design cycle	28
Figure 2	Language assessment literacy: an expanded definition	30
Figure 3	A core list of language assessment dimensions: knowledge, skills and principles	31
Figure 4	A framework for language assessment literacy development	126
Figure 5	A framework for language assessment literacy development	156



CHAPTER ONE

INTRODUCTION

The focus of this introductory chapter is on the rationale for the current thesis study discussing the problem statement. Having proposed the main purposes, research questions, significance and organization of the thesis, this chapter will serve as an overview of the following chapters.

1.1. Introduction

The national agenda of Türkiye prioritizes English language learning as an essential part of general education to remain globally competitive in terms of economy, education, technology, etc. In this sense, assessment is one of the most prominent aspects of language teaching and learning process as it ensures both accountability and sound feedback for students in addition to supporting instruction. Thus, there has been an increased emphasis on assessment practices for two decades and various forms of assessment have been discussed and proposed by prominent figures in the field of applied linguistics (Brookhart, 2011; Davies, 2008; Tsagari, 2016; Vogt & Tsagari, 2014). Given the increasing significance of assessment, teachers are expected to have a clear and full understanding of assessment types, practices and theories in order to support students' learning while providing useful feedback for teaching. As Brookhart (2011) noted good classroom assessment involves accurate inferences about students' achievement, communicating information to stakeholders, and focusing on further instruction. On the other hand, ineffective assessment causes decrease in reliability and validity, resulting in misdirected and inappropriate educational decisions. As a result, the effectiveness of assessment bears greatest importance for a number of reasons including accountability, instructional purposes, fostering growth in learning. Thus, having adequate knowledge and skills for an effective assessment is of vital importance. Given these facts, teacher assessment literacy becomes indispensable determiner for sound assessment and effective teaching practices to guide classroom teaching instructions. Therefore, nowadays there has been an increasing attention to language testing and assessment (LTA, henceforth) along with language teachers' knowledge and competence with reference to language assessment literacy (LAL, henceforth). As noted by Büyükkarcı (2016), teachers need to own necessary skills and

knowledge to interpret and use the assessment results to foster their professional improvement and student learning. More comprehensively, Davies (2008) categorizes necessary ingredients of assessment literacy under three domains as skills (know-how or basic testing mastery), knowledge (information about measurement and language), and principles (related testing concepts such as validity, reliability, and ethics). Last but not the least, one of the most essential aspects of assessment is to communicate the results to stakeholders in a meaningful way, thus maximising student learning and motivation by addressing their expectations and needs. Obviously, the basic principles of sound classroom assessment are of vital importance, otherwise improvement efforts would be ineffective (Stiggins, 1999) resulting in unreliable assessment which might be one of the biggest hindrances for students in mastering higher levels of academic excellence.

1.2. Statement of the Problem

Assessment is one of the most prominent aspects of language teaching and learning process as it ensures both accountability and sound feedback for students in addition to supporting instruction. Thus, assessment practices and various forms of assessment have gained increasing attention. More importantly, it is widely accepted that teachers play crucial role in language assessment process which requires a clear and full understanding of assessment types, practices and theories in order to support students' learning while providing constructive feedback. As Brookhart (2011) noted good classroom assessment involves accurate inferences about students' achievement, communicating information to stakeholders, and focusing on further instruction. On the other hand, ineffective assessment causes decrease in reliability and validity, resulting in misdirected and inappropriate educational decisions. In this regard, teacher candidates need to be equipped with sufficient assessment literacy comprising of language assessment concepts and knowledge in pre-service teacher education programs (Herrera & Macias, 2015). Therefore, pre-service teacher education programs are precisely of great value in order to provide adequate training in terms of LAL. However, inadequacy of teachers' assessment literacy has always been reported and discussed within the review of literature reporting that teachers do not feel confident and prepared to assess learners in an effective way as the majority of them

complained of receiving either very little or no training (Tzagari, 2016; Vogt & Tzagari, 2014).

To this end, in order to explore the language assessment needs and beliefs of stakeholders, it is considerably crucial to understand the practicing teachers and student teachers' interpretation of assessment. Further, as the underlying beliefs of individuals shape their actions; language assessment beliefs of teachers also need to be paid attention. Within this purpose, exploring LAL needs and levels along with beliefs about language assessment is worth of discussion. Considerably, it is necessary to identify the basic language assessment needs with regards to knowledge, skills and principles which might pave the way for designing an effective LTA course and LAL development training by identifying essential language assessment topics and items. As a result of this necessity, the present study attempts to identify the participants' levels, needs and beliefs about LAL in order to illuminate the contribution of pre-service teacher education program to their LAL development.

1.3. Background of the Study

Given the importance to assessment in language learning and teaching process, teachers are expected not only to manage assessment process effectively but also to interpret the data gathered from the assessment to guide and support their teaching instruction and student learning. In this sense, certain competences and qualities are undoubtedly necessary to be acquired by teachers such as selecting assessment methods, developing assessment methods, developing valid student grading procedures, communicating assessment results, and recognizing unethical, illegal and inappropriate methods of assessment which are also reported in the *Standards* (AFT, NCME, & NEA, 1990).

In an attempt to emphasize sound and efficient assessment practices in educational settings, teacher candidates need to be equipped with sufficient assessment literacy in pre-service teacher education programs. However, inadequacy of teachers' assessment literacy has always been reported and discussed within the literature showing that teachers do not feel prepared to assess their students in an effective way as the majority of them complained of receiving either very little or no training (Tzagari,2016; Vogt & Tzagari,

2014). Therefore, pre-service teacher education programs are precisely important in order to provide adequate training for LAL. However, as it was stated by Volante and Fazio (2007) teacher candidates were found with low level of self efficacy beliefs with regards to language assessment which reflects the general picture of the issue in many research studies all around the world (Chan & Jiahui, 2020; Levi & Inbar-Lourie, 2020; Tsagari & Vogt, 2017; Vogt & Tsagari, 2014). In this sense, DeLuca and Klinger (2010) acknowledged that assessment course had considerably positive impact on their assessment literacy development. Therefore, Herrera and Macias (2015) strongly advise in their studies, pre-service language education programs should include language assessment concepts and knowledge taking a more critical position. For this reason, the contribution of pre-service English language teacher education program to teacher candidates' language assessment literacy development needs to be investigated as it is estimated that the courses related to assessment might have an effect on assessment literacy of student teachers.

Remarkably, as Giraldo (2018, 2021) highlights; teachers' awareness of LAL should be raised during particularly pre-service language teacher education through language assessment courses for the following reasons:

1. Teachers are responsible for “planning, implementing, and interpreting language assessments”
2. They are the ones to communicate assessment results and make decisions about student learning
3. The general conclusion of many research studies show that language teachers are found to have limited or low levels of LAL and need further training to improve their LAL (Chung & Nam, 2018; Vogt & Tsagari, 2014).

To this end, exploring the language assessment literacy needs and assessment beliefs of in-service and pre-service language teachers is also essential in order to identify any discrepancies or overlaps between these two entities with an attempt to design a suitable training program for LAL development of teachers and teacher candidates.

However, language assessment studies have tended to lack this aspect. As a result of this gap in the field, this study attempts to identify the participants' opinions with regards to the LAL, LAL needs and language assessment beliefs in order to determine the most necessary areas for further training and LTA course content. In similar vein, Giraldo (2021) supported that an appropriate assessment based on LAL lead to a positive change and make a great deal of contribution to students learning and the program itself. Based on this perspective, as the teachers are the main implementers of the assessment activities at the practice level, their opinions and beliefs about language assessment and their language assessment literacy levels are worth of investigation.

1.4. Purpose of the Study

The purpose of this study is to explore the LAL levels and needs of EFL teacher candidates and in-service teachers, further to question language assessment beliefs and LAL development of participants. Moreover, pre-service language teacher education programs will be under scrutiny with regards to their contribution to teachers' LAL development. Considering these basic purposes, the researcher ultimately aims to make a contribution to the field of LTA by providing valuable information for the stakeholders including program developers, policy makers, teacher educators, and teachers in designing an effective and practical LTA course to improve teacher candidates' language assessment literacy knowledge, skills and competences by covering necessary topics within this course, thus preparing them for assessment practices and challenges in real classroom settings. Theoretically, the study intends to show the importance of LAL development in language teacher education as part of academic and professional success. Within this purpose, the LAL levels, needs and language assessment beliefs of in-service teachers and teacher candidates will be explored.

1.5. Research Questions

The following research questions were formulated to investigate the language assessment literacy needs and levels, in addition to the language assessment beliefs of participants:

RQ 1: What are the language assessment literacy needs and levels of pre-service language teachers?

RQ 2: What are the language assessment literacy needs and levels of in-service language teachers?

RQ3: What are the pre-service and in-service language teachers' beliefs about language assessment?

RQ4: What is the contribution of language testing and assessment (LTA) course offered in pre-service language teacher education programs?

RQ5: What are the language assessment-related topics that participants recognise as important for their language assessment literacy development?

RQ6: What are the most challenging topics in terms of language assessment literacy?

RQ7: What are the possible ways for language assessment literacy development?

As a result of pandemic (COVID-19), language testing and assessment has witnessed new trends and methods in online language assessment in addition to certain challenges and problems, that's why one more question was added to the current study as RQ8 which is formulated in the following way:

RQ8: What are the pre-service and in-service language teachers' perceptions about online language assessment?

1.6. Significance of the Study

This current research study bears importance in that it underlines certain issues with regards to language assessment literacy, language assessment training, language testing and assessment course at pre-service teacher training, and teacher beliefs of assessment. First of all, having started with the historical flow of language testing and assessment development, the study presents some background information for the following parts with

an attempt to build core information about the issue. Further, the discussion of language assessment literacy from a wide range of perspectives including the key considerations for LAL, LAL training, language assessment courses for teacher candidates, language assessment beliefs among others serves as a mirror to portray the issue from all possible angles.

What's more, the study is significant for various stakeholders such as teachers, teacher candidates, teacher educators, and program developers as it highlights the language assessment literacy needs and beliefs of EFL teachers and teacher candidates, which in turn provides necessary knowledge to design an effective language assessment course in order to equip prospective teachers with necessary competencies and knowledge by putting their needs, current levels, beliefs, and opinions for the LAL related topics at the centre. Letting the voices of pre-service EFL teacher candidates be heard provides one of the major strengths for the current study for several reasons. Firstly, they have offered crucial aspects for LAL development which might be undiscovered by teachers as test developers. Secondly, by participating in the study they have felt awareness about language testing and assessment with its core elements and significance. Moreover, as discussed by Butler (2022), other benefits of having learners in LAL discussions can be considered as a close connection to learning and instruction thus bridging the gap between teacher and learner in their understanding of assessment processes. To put it short, providing valuable information for the program developers, stakeholders, and policy makers in designing an effective and practical LTA course to improve teacher candidates' language assessment literacy knowledge, skills and competences is undoubtedly rewarding.

Further, the possible findings of this study will provide significant points for a detailed discussion of LAL training and development. This initial discussion could offer some insights on how to develop LAL of EFL teachers and teacher candidates by taking into language assessment beliefs, LAL levels and needs account. Within this purpose, a framework for LAL development was designed by the researcher based on the findings of the study which is believed as one of the major contributions of this current study to the literature. Having three main domains, the framework has 3 sub-stages each of which presents philosophical, practical and developmental aspect of language assessment literacy by discussing main considerations in every step. In this sense, this study will bear

importance as it provides a clear understanding about the contribution of not only pre-service teacher training but also language assessment beliefs to LAL development. Further, it is significant to note that online language assessment is given a prominent place within the current study by discussing challenges, coping strategies, suitable tasks and tests, and online feedback which can be considered as another contribution of this study as it provides an overall picture of the current situation after Covid-19.

Ultimately, the main discussion of this paper enables the researchers to pay attention to several significant issues, which provide fertile ground for further exploration. Initially, what might be a new research area is to explore the relationship between language assessment beliefs and LAL levels of different stakeholders. Moreover, this research highlights the need for an appropriate and efficient teacher training to enhance LAL development of prospective teachers by discussing the related language testing and assessment course in a language teacher training program for its contribution to the LAL of prospective teachers. Within this scope, the potential problems, challenges and the most effective ways for language assessment literacy development can be discussed in detail in future studies.

1.7. Context of the Study

Having adequate knowledge and skills for an effective assessment is prominent. For this reason, the contribution of pre-service English language teacher training program in general, language testing and assessment course in particular, to teacher candidates' language assessment literacy development needs to be investigated as it is estimated that the courses related to assessment might have an effect on assessment literacy of student teachers. As the purpose of this study is to explore the LAL levels of participants and further to question their beliefs about language assessment, the context of the study comprises of teaching settings and teacher education programs in general terms. One of the pre-service teacher training programs will be under scrutiny with regards to its contribution to teachers' language assessment literacy. Thus, the current study attempts to explore the contribution of language assessment course in a pre-service teacher training program on

prospective teachers' LAL levels and beliefs by demonstrating their self reported opinions about their LAL levels, needs, and language assessment beliefs. Moreover, the present study emphasizes the importance of LAL development for a language teacher as part of her / his academic and professional success so that both teaching instruction and student learning enhance simultaneously. To this end, in-service EFL teachers working at different universities in Türkiye are given a prominent place as they are the active implementers of language assessment practices in real classroom settings. Thus, portraying the issue both from pre-service EFL teachers and in-service EFL teachers' eyes provides a clear understanding of the gap between theory and practice. Discussing the gap between theory and practice with regards to language assessment in an EFL context, the present study intends to show the importance of pre-service teacher training programs for LAL development of prospective language teachers by providing valuable information for the program developers, stakeholders, and policy makers in designing an effective and practical LTA course to improve teacher candidates' language assessment literacy knowledge, skills and competences which is undoubtedly rewarding.

1.8. Limitations of the Study

Covering a mixed methods approach, this current study clearly yields valuable findings; but still it has certain limitations. Firstly, a limited number of participants cannot be representative of the EFL in-service or pre-service teacher population. Further, the data was limited to respondents' self reports; thus real life experiences and practices in classroom settings may vary. Therefore, classroom observation, field notes, reflective journals or diaries could have been useful to produce richer data. Finally, as the surveys generally attempt to show the lacking parts in teachers' assessment literacy and further the ways how to improve their understanding of assessment related topics, the respondent may have given the expected answers instead of his/her real opinions or beliefs. Though reaching a large amount of data and respondents within a rather short time; low variation in responses (Fulcher, 2012), the superficiality of answers, unreliable answers, social desirability, self-deception, acquiescence bias, fatigue effects (Dornyei & Taguchi, 2010) are among the mostly faced problems which can be also considered as limitations within the current study.

1.9. Definition of the Terms

In order to reach a clear and full understanding of the issue, it is necessary to define certain terms which would help readers orient to the topic. To start, assessment literacy can be understood in terms of competences defined in the following way by Fulcher (2012: 115):

The competences included selecting assessments, developing assessments for the classroom, administering and scoring tests, using scores to aid instructional decisions, communicating results to stakeholders, and being aware of inappropriate and unethical uses of tests (Fulcher, 2012: 115).

Concerning language assessment literacy, as a specific sub field of assessment literacy, it can be defined as covering professional knowledge and decision making skills in language education, language-related issues and language testing/assessment (Inbar-Lourie, 2008; Taylor, 2009). Further, LAL should enable EFL teachers to design appropriate assessments for particular purposes, select from a wider repertoire of assessment alternatives, critically examine the impact of standardized tests, and establish a solid connection between their language teaching approaches and assessment practices (Scarino, 2013), in addition to apprehending of key premises in second language assessment (O'Loughlin, 2006). More specifically, Davies (2008) addresses the issue by mentioning three domains including skills (know-how or basic testing mastery), knowledge (information about measurement and language), and principles (related testing concepts such as validity, reliability, and ethics), thus drawing a framework.

Due to the fact that, such terms as *beliefs, values, attitudes, judgements, opinions, ideologies, perceptions, conceptions, conceptual systems, dispositions, implicit theories, personal theories, and perspectives* are continually being uttered interchangeably (Pajares, 1992), teacher beliefs are considered to be compelling to define. But still, the following definition might shed light on the issue:

“Beliefs generally refer to those of an individual’s mental constructions that are (1) subjectively true; (2) affectively laden; (3) outcomes of substantial prior experiences; and (4) significant determiners of the individual’s actions and meaning making” (Skott, 2015: 37).

1.10. Organization of Thesis

This present study is categorized into five chapters. The first chapter serves as an introductory section which provides rationale for the study by highlighting the statement of the problem, main purposes along with research questions, and significance of the study. Chapter two documents the context of the study including historical development of language testing and assessment (LTA) together with assessment courses in pre-service language teacher education, thus drawing a theoretical framework. Having determined the theoretical framework, chapter two provides an extensive literature review which is devoted to three main sections, to name teacher beliefs, language assessment literacy, and language assessment literacy training during pre-service teacher education. Then, Chapter three illustrates the methodology retained in the study discussing both qualitative and quantitative data collection methods with a particular focus on ethical concerns and reliability-validity considerations. Moreover, Chapter four reports the findings of the study addressing related research question through a detailed discussion of qualitative and quantitative data referring to similar studies in the field. Finally, the thesis ends with salient conclusions and significant educational implications for educators, policy makers and program developers from theoretical and practical aspects along with limitations and suggestions for further research studies in Chapter five.

1.11. Chapter Summary

Chapter one has presented background information about the issue under investigation in addition to describing the purposes underpinning the study. Accordingly,

some terms have been defined in depth so that it would be easier to understand the basic concepts within the thesis. Now that an introduction to the topic has been accomplished, it is possible to move forward to Chapter 2 in order to explore theoretical framework of the thesis along with a comprehensive literature review.



CHAPTER TWO

LITERATURE REVIEW

2.1. Chapter Introduction

Having noted the underlying purposes and back ground to the study in Chapter 1, the second chapter will now address the historical development of language testing and assessment (LTA) with a specific focus on the overview of language education in Türkiye. By doing so, relevant contextual and situational factors are discussed to comprehend the theoretical framework of the thesis in order to justify the rationale of the research. For these purposes, the introductory discussion starts with a brief background information about traditional and constructivist language assessment with an attempt to probe language assessment beliefs through the prominent related research studies. The second section is devoted to language assessment literacy with an emphasis on the basic competencies and the defining qualities, further the role of pre-service language teacher training on prospective teachers' assessment literacy development will be elaborately dealt with by highlighting language assessment courses and new assessment trends like online assessment.

2.2. Theoretical Framework

2.2.1. Rationale of Research

Researchers have generally paid attention to language assessment and assessment literacy from different domains throughout the years. Concerning language assessment literacy, the international literature has generally attempted to explore LAL levels (DeLuca, 2012; DeLuca, Klinger, Pyper, & Woods, 2015; Hill, Ell, Grudnoff, & Limbrick, 2014; Koh, 2011; Koh, Carol-Ann Burke, Luke, Gong, & Tan, 2017; Lam, 2015; Mertler, 2009; Xu & Brown, 2016) and LAL needs of different stakeholders (Chung & Nam, 2018; Fulcher 2012; Hasselgreen, Carlsen, & Helness, 2004; Taylor, 2009) concluding a high need for further training as a result of the insufficiency training in language testing during pre-service teacher education (Fulcher, 2012; Giraldo, 2019; Sultana, 2019; Vogt & Tsagari, 2014). Moreover, assessment courses in pre-service teacher education programs (Kruse, Impellizeri, Witherel, & Sondergeld, 2020; Lam, 2015; Volante & Fazio, 2007)

and professional development (Koh, 2011) are among other areas of research in general sense. To start, Volante and Fazio (2007) focused on how LAL of language teachers develop through assessment courses in pre-service teacher education programs concluding that specific courses on classroom assessment were required to build high level of assessment literacy for teacher candidates. One other figure, Koh (2011) attempted to explore the contribution of professional development activities to teachers' assessment literacy while Hasselgreen, et al. (2004) and Taylor (2009) had a different purpose and focused on language assessment literacy needs which further gave path to design professional development and various training programs. More recently, Mohajer (2020) investigated the issue from a different angle questioning the relationship between assessment literacy and teaching experience among native and non-native English teachers concluding a positive relationship for native English teachers.

In Turkish context, on the other hand language assessment literacy has been investigated with similar purposes. The general tendency was on the exploration of teachers' assessment literacy levels and needs (Büyükkarcı, 2016; Kaya & Mede, 2021; Mede & Atay, 2017; Ölmezer-Öztürk & Aydın, 2018; Öz & Atay, 2017) reporting a limited LAL levels (Mede & Atay, 2017; Ölmezer-Öztürk & Aydın, 2018; Öz & Atay, 2017). However, Yastıbaş and Takkaç (2018) attempted to investigate LAL development with regards to instructional purposes and classroom implications through a think aloud protocol with 8 participants, revealing that participants were mostly found as course book and student oriented while developing exams.

When it comes to teacher beliefs, beliefs about the purposes of assessment and the practices have been probed by Barnes, Fives, and Dacey (2017), Opre (2015) and Thomas (2012). On the other hand, in the area of language studies, studies conducted by Cheng, Rogers and Hu, (2004); Muñoz, Palacio, and Escobar (2012); Rogers, Cheng, and Hu, (2007), and Wicking (2017) focused on language assessment beliefs and the way they shape their teaching and assessing behaviours while Chang (2006) particularly investigated oral language assessment practices. From a different point of view, Hakim (2015) studied the factors affecting the teachers' beliefs about assessment concluding that teaching time

and curriculum workload negatively influenced teachers' assessment practices and beliefs about the use of different assessment types. Considering Turkish context, teacher beliefs about assessment have been paid attention by various fields of study such as mathematics (Aydın, Baki, Köğce, & Yıldız, 2009); language particularly English (Büyükkarcı, 2014; Öz& Atay, 2017; Şahinkarakaş, 2012). The general tendency of researchers was on the assessment perceptions and practices of teachers within different contexts including primary education, tertiary level, teacher education programs, and EFL settings.

More importantly, the needs, beliefs, levels and perceptions of assessment may vary among different stakeholders (Taylor, 2013; Yan & Fan, 2020), thus each stakeholder including language teachers, student teachers, teacher educators, language testers, program developers and policy makers should be examined in detail to gain rich insights. To my knowledge, based on the brief literature review above although research has analysed language assessment courses and its impact on teachers assessment literacy, but still identifying certain areas of both pre-service and in-service EFL teacher language assessment literacy needs and their beliefs about language assessment and practices, is still underexplored. Therefore, research into how language assessment literacy develops during pre-service teacher education and / or teaching experience and how assessment beliefs and language assessment literacy interact seem very useful, due to the fact that teacher beliefs somehow shape the way of teachers' teaching instruction and assessment. Thus, research still needs to consider an EFL setting to see the basic competencies language teachers need to have for a sound and effective assessment in addition to explore language assessment beliefs. Therefore, this study seeks to contribute to the field of language assessment by providing some insights into how prospective and practicing EFL language teachers enhance their LAL in their pre-service language teacher education in Türkiye by analyzing their perceived LAL levels, needs and language assessment courses they are offered in addition to exploring their language assessment beliefs. In this sense, the following part is devoted to discussing historical development of LTA to provide background knowledge of the issue under investigation.

2.3. Historical Background

In order to reach a comprehensive understanding of the language testing and assessment, it is quite necessary to iterate the historical development of LTA from various domains such as epistemological, methodological and theoretical aspects. Because the philosophical perspectives behind a trend or an approach bear the greatest importance for justifying the underlying reasons and purposes, the shift from one point of view with regards to LTA to another can be only understood clearly by presenting the basic philosophical tenets behind them.

In this regard, the evolvement of LTA might be comprehended distinctly in relation to the developments in language teaching and learning approaches. Traditionally the field of LTA has been debated and constructed based on the underlying views about teaching and learning language, thus the two might be regarded as being interrelated. Consequently, when the language teaching methodology has changed, not only the purposes of testing but also testing strategies and techniques have gone through certain changes. As a result, not only how to test but also why to test and what to test have changed dramatically which results in differences in the test methods and the interpretation of the test results accordingly. For this reason, the periods in the history of LTA will be discussed by crystallizing the corresponding language teaching and learning approaches. Considering the evolvement of LTA, Spolsky (1978) declared three basic stages, named as the pre-scientific, the psychometric-structuralist, and the integrative-sociometric. In line with this, the historical development of LTA is going to be dealt with under these perspectives. In doing so, the basic tenets of each approach and the shifting from one approach to the other can be pictured vividly.

To initiate, prior to 1960s, in pre-scientific stage, the basic purpose of language learning was to acquire the target language literature, thus having an extensive knowledge of vocabulary and grammar rules (Giri, 2003). At the core of pre-scientific period is grammar translation method in which students are expected to translate the text either into or from the target language (Farhady, 1979). Accordingly, the language testing aimed at

identifying the accuracy of vocabulary and the grammatical forms and structures within written papers. These kinds of tests were claimed to be time saving and simple to construct. However, as a shortcoming of this trend can be stated as a lack of reliability, objectivity, context and functional use of language (Giri, 2003; Mahapatra, 2019).

Secondly, the 1960s witnessed a new approach to language testing- the psychometric-structuralist period- based on the views of both behaviouristic learning theory which supported habit formation and structural linguistics focusing on the idea of language being segmented into items. According to this approach of language testing, statistically sound tests were accepted as valid and reliable in which each item of language (vocabulary, reading, grammar, pronunciation, etc.) is tested separately. This way of testing is also known as discrete-point test which supports the idea that language can be segmented into its elements, such as grammar, vocabulary, reading, pronunciation, etc. In these tests, each question attempts to test only one aspect of language separately, thus they have been criticized harshly as they do not assume language as a whole with a certain context in which the language is used in addition to lacking individual learner abilities and functional aspect of language within a real life communication (Mahapatra, 2019).

As a third period -psycholinguistic/sociolinguistic language testing- also known as integrative approach to language testing supported that language cannot be isolated from culture and social context, thus suggesting an actual use of whole language. In this period, Carroll (1961) supported the need for testing communicative abilities of language learners within a language test. In association with this trend, language testing was based on the communicative use of language in social context, shifting “its emphasis from linguistic accuracy to functional ability” (Giri, 2003: 59), thus appropriateness took the first seat. “[C]loze test, dictation, listening and reading comprehension, and oral interviews” (Farhady, 1979: 348) are among the integrative tests. More importantly, after Hymes (1972) emphasized communicative language competence model, the language testing had a significant focus on both communicative and sociolinguistic aspects of language. Accordingly, Canale and Swain (1980) and Bachman (1990) proposed communicative language testing in which appropriateness, actual language use in a meaningful context

were among the major purposes. Further, Bachman and Palmer (1996: 27) suggested that communicative language testing consisted of four dimensions, including “specificity of context, authenticity of materials, authenticity of test tasks, and simulation of real life situation” with a particular emphasis on four types of competencies -linguistic, sociolinguistic, discourse, and strategic. When the meaning and context are concerned, it is also worth mentioning another approach which is called pragmatic language testing which has received increasing attention with a more precise focus on meaning and specific context (Oller, 1978 cited in Farhady, 1979).

More recently, taking a critical position to testing has gained more popularity and importance discussing its role in terms of political and ideological domains. Inspired by Paulo Freire’s (1970) masterpiece ‘Pedagogy of the Oppressed’, applied linguists have taken seriously the idea of language testing having a hidden agenda with certain biases. In this sense, significant figures in language testing like Spolsky, Tim McNamara, Elana Shohamy accepted language testing as a “powerful educational tool that is used for social, political, cultural, and above all, ideological control” (Mahapatra, 2019: 8). Therefore, new critical concepts “*bias, hegemony, democracy, marginalization, dominance, ideology, fairness, ethics etc*” occupied crucial place in *critical language testing* (Shohamy, 2001). In doing so, critical language testing attempts to take individual learner differences, and identity of learners into account.

To put it short, assessment in the process of teaching and learning language has widely changed in accordance with the change of the attitude towards teaching and learning language from behaviourism-oriented to constructivist-based assessment (Rueda & Garcia, 1996). That’s why this background knowledge about the development of language assessment is beneficial to reiterate the evolvement of LTA. So far, the historical development of LTA is briefly mapped out at epistemological, methodological and theoretical level. Having covered the basic periods of language testing and development, a comparison of traditional language assessment and social constructivist language assessment will now be provided.

2.3.1. Traditional Language Assessment versus Social Constructivist

Defining what learning is leads one to define what assessment is and what assessment scores mean accordingly. In this sense, it is quite reasonable that both traditional understanding of assessment and constructivist assessment view learning in a different way. To start, traditional definition of learning is regarded as rule learning, and assessment shows whether the rule is learnt or not, hence being product oriented (Delandshere & Jones, 1999). In accordance with the traditional approach to language assessment, lower-order thinking skills are measured through true-false, matching, multiple-choice, gap-filling and short answer methods (McMillan, 2014; Popham, 2014; Tao, 2014). As Popham (2014) reported traditional assessment has certain advantages in terms of administration and scoring. However, its being narrow sided involving lower level thinking skills causes a lack of meaningful learning without a real life context (Miller, Linn, & Gronlund, 2013).

On the contrary, constructivist assessment views learning as process oriented which suggests that learning is a continuous development of learners. Accordingly, assessment plays a facilitating role providing constructive feedback (Ogan-Bekiroglu & Suzuk, 2014). Moreover, constructivist language assessment with a greatest focus on higher-order thinking skills includes performance-based assessment, alternative assessment such as self assessment, peer assessment and portfolio assessment methods (Brown & Hudson, 1998; McMillan, 2014; Miller et al., 2013; Russell & Airasian, 2012). In contrast to traditional assessment, constructivist language assessment provides a meaningful learning environment for learners as it fosters self and peer assessment, portfolio assessment which include innovative assessment tools such as “assignments, interviews, problem-solving tasks, communicative pair-work, role playing, observations, journal writing and group discussions” among others (Tao, 2014: 37). Even though constructivist assessment, also known as innovative or alternative assessment enhances lifelong learning, it still has faced a number of criticism in terms of scoring (inter-reliability and/or intra-reliability) and sample tasks (construct validity, content validity, to name a few), let alone its being time consuming (McMillan, 2014; Popham, 2014). Constructivist assessment aims at promoting learning rather than scoring and ranking learners. According to this

perspective, students are accepted as active participants in assessment process as they supervise their own learning, and contribute to their peers' learning by giving feedback (Inbar-Lourie, 2008a). As Filer (2000) stated constructivist assessment, based on contextual and situational understanding, views assessment as “*a social practice and a social product*” which yields certain consequences for particular groups or individuals.

Moving the discussion further, not only the underlying epistemological understanding of assessment but also the purpose and the impact of assessment shape the idea of assessment which can be regarded as *assessment of learning, assessment for learning, and assessment as learning* approaches. Initially, assessment of learning supported by traditional domain is considered as summative assessment which mainly focuses on the outcome at the end of a unit or a whole term based on some criteria or standards. Secondly, assessment for learning, also known as formative assessment ensures an interaction between learner and teacher based on the feedback with an aim to enhance learning. Assessment as learning, on the other hand, assists learners in monitoring their own learning and through the feedback they gather from the monitoring they make necessary changes and adaptations (Volante & Fazio, 2007).

Having portrayed a general overview about the historical development of language assessment with a specific focus on traditional and constructivist assessment approaches, teacher beliefs related to assessment and language assessment in particular will be discussed in the following part.

2.4. Teacher Beliefs

Research into the correspondence between the language assessment literacy and teacher beliefs is quite prominent as teacher behaviours in assessment process are shaped by their beliefs, thus ensuring a deep understanding of LAL development. Within this view, the following part aims at highlighting the teacher beliefs in a detailed manner.

Teacher beliefs, consciously or unconsciously, shape their understanding, behaviours, and classroom practices. These beliefs about learners, learning, instruction, and

assessment provide foundation for their classroom practices (Thomas, 2012). Because of its complicated and abstract nature, *belief* is hard to define and analyze accurately. But still, various definitions and terms might be seen in the literature as teacher ‘assumptions, perceptions, conceptions, and values’ which affect their actions, goals, knowledge, decision making processes in general terms, thus having a considerable amount of impact over teacher’s teaching instructions and interpretations (Barnes, Fives, & Dacey, 2015; Brown, 2006; Irez, 2007; Matese, Griesdorn, & Edelson, 2002; Opre, 2015; Rogers, Cheng, & Hu, 2007; Thomas, 2012). Notably, Skott makes a comprehensive definition of beliefs as:

“an individual’s mental constructions that are (1) subjectively true; (2) affectively laden; (3) outcomes of substantial prior experiences; and (4) significant determiners of the individual’s actions and meaning making” (2015: 37).

Taking a more social and critical perspective, Wicking (2017) noted that assessment, teaching, and learning processes are considered to be cultural, historical, social, and context oriented. In this regard, assessment from socio-cultural perspective is assumed as context oriented which is carried out in social settings (the local school environment). This point of view is based on Vygotsky’s socio-cultural theory which suggests that learning is not independent from its social, historical and cultural context. As Chang (2006: 3) stated in his study, assessment within socio-cultural theory is considered as “*interactive, dynamic, collaborative, constructive, and enabling*”. In this sense, teacher beliefs about language learning and assessment can be considered as a strong indicator of their roles and behaviours in classroom setting as a social context. Remarkably, Chang (2006) stressed the importance of teachers’ experiences as learners and their interaction with social and cultural context in shaping their assessment beliefs.

Concerning language assessment, the way teachers perceive assessment shapes the way they use assessment types and tools and the way they interpret assessment results (Brown, 2008; Wicking, 2017). Therefore, as Brown (2008) discussed, teacher beliefs about the purpose of assessment play a major role in their assessment practices. Moreover,

these beliefs are considered to have an effect on the way students are engaged in assessment processes (Barnes, et al. 2015, 2017). In short, not only the purpose but also the selection, development of assessment and interpreting of the assessment results are influenced by teacher beliefs. In terms of assessment purposes, Barnes, et al. (2017: 108) formulated the beliefs “on a continuum from an extreme pedagogical to an extreme accounting (accountability) conception”. Similarly, Brown, Chaudhry, and Dhamija (2015: 109) framed a four factor model in which the purposes were summarized as “assessment for improvement, assessment as irrelevant, assessment as a tool to control lessons and teaching, and assessment as an indicator of school quality” (cited in Barnes, et al. 2017).

When it comes to the relevant research studies, a study conducted by Lu (2003 cited in Chang, 2006) through interviews and classroom observations revealed a high consistency between beliefs and assessment practices. Through an ethnographic study, Chang (2006) attempted to explore instructors’ assessment practices and beliefs. He found that the two instructors had both objective and subjective tools to assess students taking into account students’ academic work and non academic work (effort, participation, etc.) as well. In his study (2006: 14), he noted six themes about teachers’ beliefs with regards to oral language assessment as follows: 1) language learning, 2) orientations and purposes of assessment practices, 3) decision making rationale, 4) grading practice, 5) use of the result gained from assessment, 6) their roles in oral language assessment concluding that the purposes of assessment could be categorized under three constructs: student-centred, instruction-based, and administration-based assessment (2006: 18). He stressed that what the teachers believed about the purposes of assessment influenced their assessment methods and practices. In terms of decision making process, teachers reported three main factors affecting their beliefs as (a) the nature of learning objectives, (b) the importance of using a wide range of practices, (c) professional experience (Chang, 2006: 21). Concerning the assessment results, teachers were found to use assessment results to alter both their practices and students’ actions, thus supervising and enhancing students’ learning.

In similar vein, Rogers, Cheng and Hu (2007) attempted to explore teacher beliefs about assessment in Canada, Beijing and Hong Kong through a questionnaire discussing

assessment procedures, the value of assessment, and the time needed to engage in assessment, and the place of standardized testing. They also grouped assessment purposes as assessment for learning, assessment as learning, and assessment of learning. In their study, English teachers in Japan were found to be more interested in assessment of learning because of the robust influence of standardized language testing embedded there. Determining students' final grades was the most frequently purpose stated by the teachers ($n= 134$) while preparing students for a standardized test was the least ($n= 27$). In short, the participants in three contexts (Canada, Beijing, in Hong Kong) were found using assessment for student centred purposes by employing formative assessment through feedback, and summative assessment for giving final marks.

What's more, Wicking's study (2017) showed that teachers mostly used performance based assessment. More saliently, native speaker teachers were found to believe that assessing speaking and listening was better for assessing learning than paper and pencil assessments than non-native speaker teachers. In general sense, teachers reported using assessment to give feedback to students rather than to use assessment results for their enhancement and professional development.

As an extensive overview, Matese, et al. (2002: 6) presented a literature review in which they put assessment beliefs in three groups such as "*beliefs about the goals and nature of assessment, beliefs about assessment constraints, and perceptions of assessment self-confidence*". The first category suggests assessment being formative or summative, subjective or objective; while the second category addresses some problems like the time limit. The last part is related with teachers themselves and their assessment practices.

While Matese, et al (2002) categorize assessment beliefs in terms of purposes, constraints, and self-confidence; Thomas (2012: 106) identified and reported seven beliefs of teachers about assessment methods, tools, and purposes as in the following:

- (1) Assessments which take place informally in the class are the best ways of assessing students' performance,

- (2) Informal assessments are a waste of teaching time,
- (3) Assessment is a joint venture between teachers and parents,
- (4) Assessment encourages students to look critically their own classroom performances,
- (5) Assessments in the form of direct observation, reduces students' academic achievements,
- (6) Assessment pressurizes teachers to complete their syllabi,
- (7) Assessment, in the form of formal tests, makes a negligible contribution to student learning (2012: 106).

More comprehensively, Opre (2015: 231) pointed out four basic teachers' beliefs which were deducted from the related literature in her study as follows:

- a) [A]ssessment improves teaching and learning;
- b) assessment holds the students responsible for their own learning;
- c) assessment charges institutions and teachers with responsibility of students/pupils teaching;
- d) assessment is irrelevant, it negatively affects teachers, students/pupils, curricula and teaching (2015: 231).

The first type of teacher beliefs highlights the role of assessment in enhancing students' learning through supportive feedback and improving the teaching. The second view supports the idea that students need to manage their own learning while the third view puts the responsibility on teachers and institutions concerning the breakdowns in students' learning. Lastly, the negative effects of assessment on teachers, students, and teaching are addressed. However, on the contrary to what the previous studies reported, a clear divergence can be witnessed between teacher beliefs and their practices in Muñoz, et al (2012). In their study, the teachers self reported that they believed in assessment for improvement; however, they were found to use summative assessment more frequently.

Concerning Turkish context, Şahinkarakaş (2012) conducted a study to explore teacher beliefs about language assessment and the impact of teaching experience on their beliefs, if any. In his study, he employed a qualitative research design in which the participants were asked to use a metaphor instead of language assessment. Through theme based coding, he reached 73 metaphors, 27 of them implied formative assessment, for ex. *“lifelong learning, deep-sea diving, checking a roadmap, a box ready to be opened, a window with a lot of angles”* (2012: 1789). On the other hand, the metaphors addressing summative assessment were indicated as *“end of a journey, harvesting the crop, a finish line of a marathon, the medal received at the end of a race”* (2012: 1789). When it comes to teacher assessment beliefs and their classroom practices, some studies seem notable (Büyükkarcı, 2014; Han & Kaya, 2014; Şahin, 2018). While formative assessment was mostly supported on the basis of participant responses (Büyükkarcı, 2014) but still summative assessment took the first place in practice (Şahin, 2018), thus a misalignment between teachers’ beliefs and practices can be observed.

At its simplest, belief has been the major focus of research studies in understanding teachers’ thoughts, practices and changes in practices (Irez, 2007). But more importantly, teachers can further improve their teaching skills when they are aware of their beliefs and the impact of beliefs on their teaching and learners (Freeman & Johnson, 1998). Although the impact of teacher beliefs on their language assessment processes and practices has been well established (Brown, 2008; Opre, 2015; Rogers, et al., 2007), the role of teacher beliefs in language assessment literacy has not been investigated. Teachers’ beliefs about assessment inform their thinking and planning, consequently shape their assessment practices. Because of the fact that teacher beliefs lead one to reach a clear understanding about teachers’ classroom practices, it is quite prominent to investigate their beliefs to gain insight into their language assessment literacy development in an EFL context. Therefore, in this study, teachers’ beliefs about language assessment and language assessment literacy will be under scrutiny as assessment is one of the most important decision making processes in which collecting and recording information about students’ learning and communicating the results are among the main purposes.

Having presented an extensive introductory part about teacher beliefs, now it is quite remarkable to embody language assessment literacy by throwing shed light on relevant literature.

2.5. Language Assessment Literacy

Assessment is one of the most prominent aspects of language teaching and learning process as it ensures both accountability and sound feedback for students in addition to supporting instruction. In this regard, Xu and Brown supported the idea that assessment literacy is an important integral part of teacher expertise (2016). Thus, positioned at the heart of education, teacher assessment literacy shows the overall quality and accountability of education. Within this scope, there is an increased emphasis on assessment practices and various forms of assessment.

Given the increasing significance of assessment, language teachers are expected to have a clear and full understanding of assessment types, practices and theories in order to support students' learning while providing useful feedback for teaching. For this reason, language teacher education programs bear prominent importance as they equip prospective teachers with necessary skills, knowledge, and practices with regards to language assessment. In similar vein, Herrera and Macias (2015) strongly advise in their studies that pre-service language teacher training programs should include language assessment concepts and knowledge taking a more critical position. However, inadequacy of teachers' assessment literacy has always been reported and discussed within the literature showing that teachers do not feel prepared to assess their students in an effective way as the majority of them complained of receiving either very little or no training (Tzagari, 2016; Vogt & Tzagari, 2014). With an aim to fill this gap, this research study attempts to explore the language assessment levels and needs of in-service EFL teachers and pre-service student teachers in order to understand their perceptions and interpretation of language assessment, further their language assessment beliefs and language assessment literacy (LAL, henceforth) development are under investigation in this study.

Initially, to gain a clear understanding of the term, a definition could be helpful. In this sense, assessment literacy can be defined comprehensively in terms of the following competencies which were emphasized by Fulcher (2012: 115):

“selecting and developing assessments for the classroom, administering and scoring tests, using scores to aid instructional decisions, communicating results to stakeholders, and being aware of inappropriate and unethical uses of tests” (2012: 115).

In general terms, the *Standards* (AFT, NCME, & NEA, 1990) prescribe seven competency domains in which teachers should be skilled; that is,

1. Choosing assessment methods appropriate to instructional decisions;
2. Developing assessment methods appropriate to instructional decisions;
3. Administering, scoring, and interpreting the results of both externally produced and teacher-produced assessment methods;
4. Using assessment results when making decisions about individual students, planning teaching, developing curriculum, and school improvement;
5. Developing valid pupil grading procedures;
6. Communicating assessment results to various stakeholders; and
7. Recognizing unethical, illegal, and inappropriate assessment methods and uses of assessment information.

These competences cover general assessment knowledge, skills and principles in a broad sense. Through the standards and competences, language teachers are expected to reach a deep understanding of language assessment and testing procedure. As the test design and development is a complex and multifaceted process in which there are various

stages from the development of test items to interpreting the scores and communicating the results, what is prominent to recognise is that it has to be precisely planned in detail. In this regard, the following figure designed by Fulcher (2010) is a comprehensive one to explain the basic steps in test design and development.

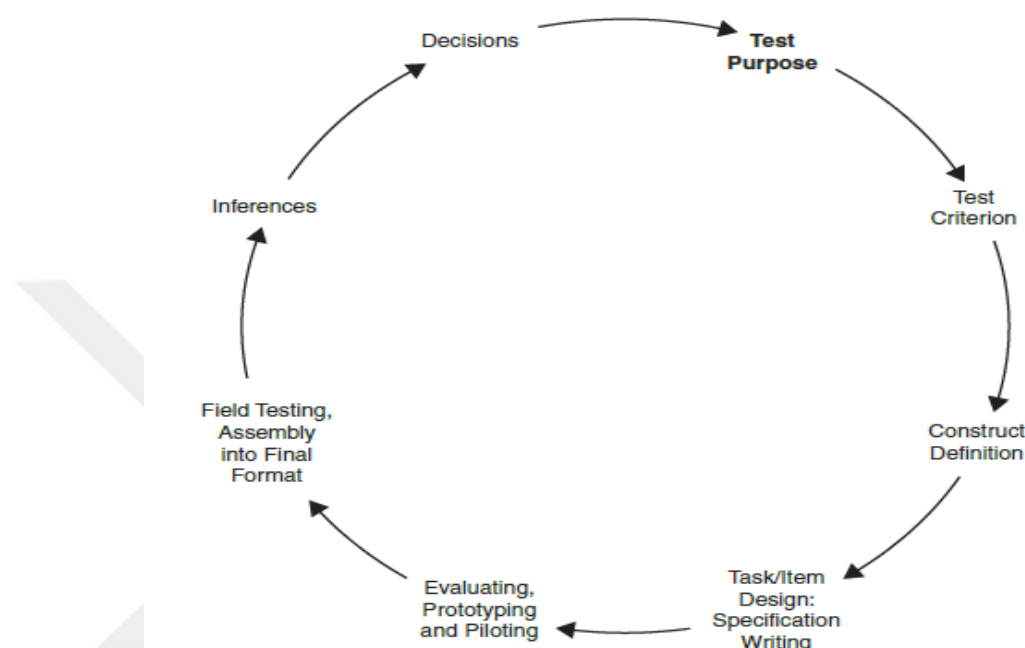


Figure 1: The test design cycle (Fulcher, 2010: 94).

Initially, according to the Figure 1, the process of test design starts with clear and necessary decisions for a purpose (why to test), for the content (what to test) and for the method (how to test) which then shapes the test specifications on which the test is built. Moreover, what to test is decided by the course content and pace of the course while how to test and why to test are differentiated on the basis of teachers' assessment beliefs or administrative decisions. In each step, you need to identify the rationale behind your choices which are fed by your underlying beliefs and prior experience. Writing test specifications is another LAL item regarded as important for fair and meaningful assessment which provides a detailed plan for the development of a test. Scoring an item is another important tenet for the process of developing a language test as put forward by Fulcher in his own words: "if it is unclear how an item is to be scored, or if credit is being

denied or given for construct-irrelevant reasons, the item is not going to provide useful information that contributes towards the meaning of the test score” (2010: 201). While designing a good and sound language test requires a great amount of time, effort, knowledge and experience with an attempt to produce meaningful and fair results for learners; selecting ready-made tests, on the other hand is considered easy to do. However, it is unfair to say so because selecting ready-made tests and using them in your own context also necessitates to consider a certain number of factors involving test purpose, individual differences of learners, test content, practicality and administration of the test, washback effect of the test, reliability and validity issues among others (Fulcher, 2010).

In synthesis, language assessment literacy requires not only knowledge about testing but also certain related skills which will be detailed in the next section.

2.5.1. Defining Qualities of Language Assessment Literacy

This part is going to discuss the definition of language assessment literacy and the basic underpinning traits teachers need to have. Defining LAL is regarded as challenging because it is notably difficult to decide what being assessment literate requires (the content), how it can be developed (skills), and why it is important (rationale). For two reasons put forward by Giraldo (2021), there is no exact definition for LAL. While the first reason is concerned with changing roles and profiles of various stakeholders such as language testers, teachers, administrators; the second one shows the lack of a specific authority to determine the core features of LAL.

Although being interchangeable in nature, still some significant figures have offered various definitions for language assessment literacy emphasizing different aspects of language. To start, Inbar Lourie (2008a: 389) defines assessment literacy as "having the capacity to ask and answer critical questions about the purpose for assessment, about the fitness of the tool being used, about testing conditions, and about what is going to happen on the basis of the results” while emphasizing LAL as the ability to comprehend, analyze and use assessment data with an attempt to improve students’ learning and their instruction (Inbar-Lourie, 2008). Therefore, LAL includes language teachers’ familiarity with the purpose of assessment, various ways of test development, and using assessment results

appropriately. Weideman (2019), on the other hand discussed the term with three basic terms as consciousness, awareness and knowledge of assessment in his definition in which he highlights the interpretability of the scores or communication the results to the learners to make a meaningful assessment. However, having adequate knowledge about LAL is not meaningful unless it is armed by context sensitive knowledge and situational understanding of the context, thus conceptualized as more complex and dynamic by Scarino (2013).

Although not an explicit definition is offered for LAL, some models are presented to identify specific qualities and competences for LAL. Thus, dealing with LAL as a conceptual framework, it is necessary to mention the following models. Firstly, Brindley (2001) proposed “what, how, and why” questions with regards to LAL by discussing the theoretical background (what), methods (how), and reasons (why) for language assessment (Inbar-Lourie, 2008; Scarino, 2013). Similarly, Davies (2008) designed a framework including “*skills, knowledge, and principles*” and argued that in order to be assessment literate, necessary skills such as item writing, statistics, test analysis, etc., relevant background knowledge about language learning, teaching, and testing, furthermore some principles such as ethics, fairness, etc. are needed.

Another three dimensional model was presented by Fulcher (2012) who suggested “contexts, principles, and practices” as key competences for assessment literacy portrayed in Figure 2 below.

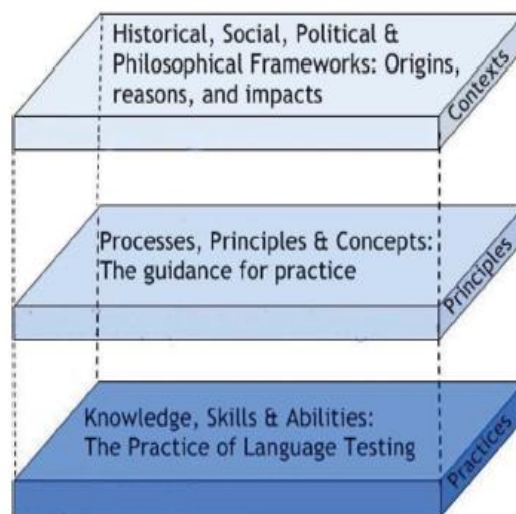


Figure 2. Language assessment literacy: an expanded definition (Fulcher, 2012: 126)

As it is shown in the Figure 2 above, “contexts” addresses historical, social and political considerations while “principles” suggests the general processes and concepts for practice of assessment. Moreover, “practices” are concerned with knowledge, skills, and abilities for assessment. Considering theory and practice of language assessment, Vogt and Tsagari (2014: 377) highlighted core competences of LAL as “the ability to design, develop, and critically evaluate tests and other assessment procedures, as well as the ability to monitor, evaluate, grade and score assessments on the basis of theoretical knowledge”.

To draw an overall analysis, Giraldo (2018) designed a frame based on an extensive literature review discussing the core list of LAL dimensions named as knowledge, skills and principles by Davies (2008) as displayed in the following Figure 3.

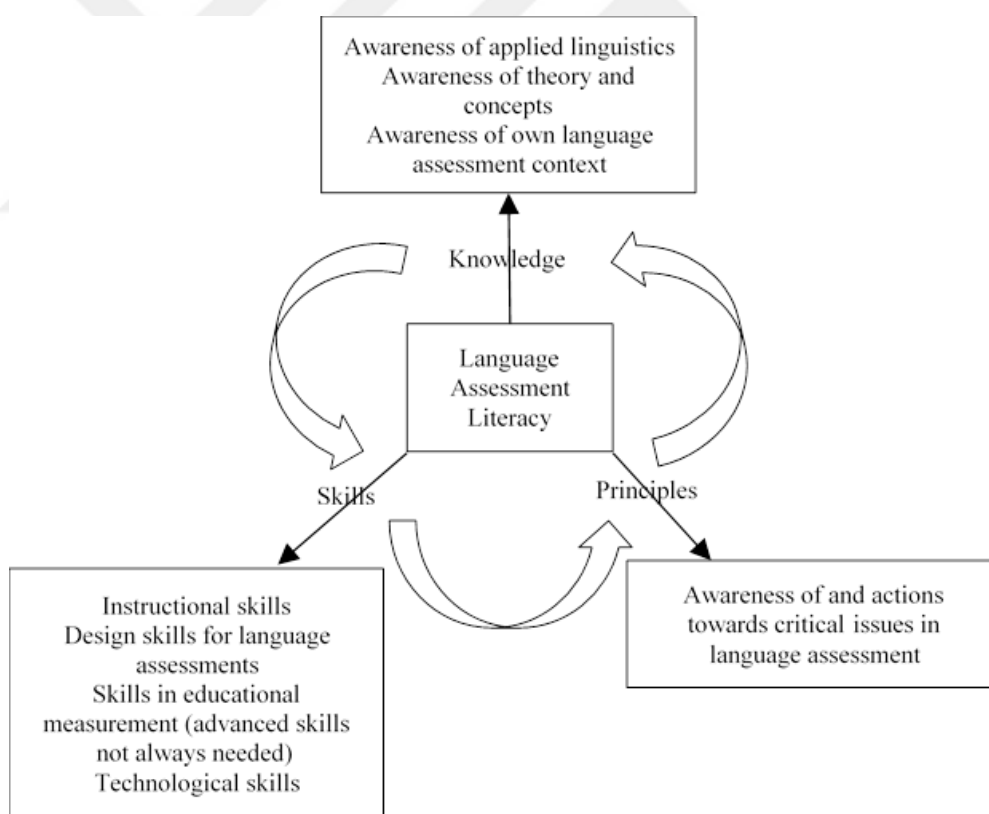


Figure 3. A core list of language assessment literacy dimensions: knowledge, skills, and principles (Giraldo, 2018: 187).

Based on the model of Davies (2008), Giraldo (2018, 2021) discussed the phenomenon from two perspectives to name development perspective which is concerned with a process from planning to evaluating language assessments; and knowledge perspective which mostly emphasizes understanding of the results and making decisions. In particular, *knowledge* concerns with awareness of language related theory and concepts while *skills* cover instructional, design and technological skills for language assessments. The last domain is language assessment *principles*, including awareness of critical issues in language assessment.

On the contrary to the aforementioned thoughts and models by significant figures in the literature, Baker and Riches (2018) drew attention to socio cultural aspect of language assessment underlining collaboration and self awareness which in turn paves the way to “contextualized, collaborative, and reflective nature of LAL”. Yan and Fan (2020) noted LAL is accepted as “social and co-constructed”, therefore accumulating a core list of knowledge, skills and principles is no longer adequate for an effective LAL, but rather a context sensitive, self reflective and experience based LAL is supported (Crusan, Plakans, & Gebiril, 2016; Scarino, 2013; Yan, Zhang, & Fan, 2018). Finally and more comprehensively, Taylor (2013: 410) paid attention to differing needs and levels of various stakeholders in terms of LAL, thus offered another model consisting of eight dimensions as:

1) knowledge of theory, 2) scores and decision making, 3) technical skills, 4) principles and concepts, 5) language pedagogy, 6) personal beliefs/attitudes, 7) local practices, and 8) socio-cultural values (2013: 410).

That’s why in this current dissertation, the researcher has planned to obtain qualitative data in order to display the reflections and experiences of different stakeholders within their contexts from their own eyes, thus letting their voices be heard.

2.5.2. Relevant Research Studies

Three leading tendencies could be noted for LAL research foci and discussions. The first focuses on the basic skills, knowledge and competences teachers are expected to have (Fulcher, 2012; Giraldo, 2018; Inbar-Lourie, 2013) while the second research and discussion emphasize the needs of pre-service and in-service language teachers (Fulcher, 2012; Giraldo & Murcia, 2018; Vogt & Tsagari, 2014). Additionally, the third research focus mostly deals with teachers' LAL development (Baker & Riches, 2018).

Moving the discussion further, the vast of research studies seemed to be designed as deficiency oriented, because most of them aimed at exploring the current levels of LAL of various stakeholders -generally assumed as low- (Chan & Jiahui, 2020; Levi & Inbar-Lourie, 2020), the assessment related areas that need more training as a result of insufficient teacher training (Tsagari & Vogt, 2017; Vogt & Tsagari, 2014), the contribution of assessment courses (Jeong, 2013; Jin, 2010; Kruse, et al 2020; Sevimel Şahin, 2019). More specifically, much of the LAL studies in the literature to date has focused on the current LAL levels of language teachers (DeLuca, 2012; DeLuca et al., 2015; Hill et al., 2014; Koh, 2011; Koh, Burke, Luke, Gong, & Tan, 2018; Lam, 2015; Mertler, 2009; Xu & Brown, 2016), LAL related needs (Chung & Nam, 2018; Fulcher 2012; Hasselgreen, Carlsen, & Helness, 2004; Taylor, 2009; Vogt et al. 2008), language assessment courses (Brown, 2008; Kruse, et al 2020; Lam, 2015; Volante & Fazio, 2007), language assessment literacy development through professional development (Fulcher, 2012; Koh, 2011). The most salient results from relevant research studies are going to be summarized and discussed below.

The first eye catching result of several studies can be stated as low level of language assessment literacy of language teachers (Volante & Fazio, 2007; Xu & Brown, 2016). As reported by the studies above, many language teachers feel unprepared to assess their learners due to inadequate assessment training in pre-service teacher education programs. Secondly, several studies (Chung & Nam, 2018; Hasselgreen, Carlsen & Helness, 2004; Vogt & Tsagari, 2014) have revealed that teacher training programs do not provide adequate assessment literacy training for prospective teachers. Thirdly, concerned

with the language assessment courses, Bailey and Brown (1995) examined various course characteristics, including topics covered and core textbooks used while O'Loughlin (2006) explored the course through an online forum with the written comments of two participants.

To illustrate, Volante and Fazio (2007) carrying their study with 69 teacher candidates in Canada reported low level of self efficacy beliefs with regards to language assessment while DeLuca and Klinger (2010) administered another study with 288 teacher candidates concluding that assessment course had considerably positive impact on their assessment literacy development. Concerning the differences between pre-service and in-service teachers, Mertler (2003) found that in-service teacher assessment literacy levels were significantly higher than the pre-service teachers' levels, which shows that classroom experience has a positive impact. With regards to teacher training programs, Hasselgreen, Carlsen and Helness (2004) examined the assessment training needs of teachers in Europe and reported that the training did not fully prepare participants to perform assessment-related activities. Similarly, Vogt and Tzagari (2014) reported insufficiency of assessment training in their study by using the same survey with 878 foreign language teachers. Moreover, Chung and Nam (2018) carried out their study in Korea and showed a lack of training in LAL and the participants' discontent with in-service training in language assessment.

From a socio cultural point of view, some studies highlighted "contextualized, collaborative, and reflective" nature of LAL. For example, some significant figures supported contextual, experiential concerns (Yan & Fan 2020), self-reflection and apprenticeship (Scarino, 2013), collaborative and reflective assessment practices (Harding & Kremmel, 2016; Levi & Inbar-Lourie, 2020; Vogt & Tzagari, 2014) for LAL development. Yan and Fan (2020) pointed out that language teachers can improve their LAL through collaborative assessment activities and communication with language testing researchers, assessment coordinators, teacher mentors, or experienced instructors in an apprenticeship-based, experience-mediated model. As previously noted, both contextual and experiential factors shape assessment practices of language teachers by influencing

their tendency to use a certain type or tool of an assessment as they mostly prefer to use familiar assessment methods they have experienced as test takers (Reynolds-Keefers, 2010). Thus, contextual aspect of assessment represents a wide range of parameters from educational to historical and social life while experiential aspect, as the name suggests, is related with the individual's assessment background involving training, real life practices, beliefs and one's own experiences of assessment. With a different group of stakeholders, Kremmel and Harding (2020) tried to explore language assessment literacy needs of teachers, language testers and language testing researchers in an empirical research reporting significant differences across stakeholders. Lam (2019) on the other hand drew attention to a different aspect of assessment by investigating writing assessment and knowledge of particular practices of secondary school teachers in Hong Kong reporting an adequate knowledge level of participants and positive beliefs for alternative writing assessments. However, the studies investigating LAL of learners are scant. For instance, Butler, Peng, and Lee (2021) explored the LAL of young learners in China reporting that the children in the study had knowledgeable views about assessment as they drew attention to the lack of communicative tasks in their assessments.

At the national level, LAL studies mainly focus on two main group of stakeholders, to name in-service teachers' LAL levels, LAL development and LAL training needs (Hatipoğlu, 2010; Sevime Şahin, 2019; Yastibas & Takkac, 2018; Yetkin, 2015; Yastibaş, 2018), and pre-service teachers' LAL levels, LAL development and LAL needs (Kömür, 2018; Şahin, 2018). Considering Turkish context, even though the main foci of research studies do not change, it is still very significant to mention some basic studies with their results to reach a better understanding of contextual knowledge for this current study itself. Initially, a mixed methods research study done by Mede and Atay (2017) through questionnaire and focus group interviews revealed limited level of LAL of Turkish EFL teachers particularly discussing deficiencies in test design and procedures in addition to specific terms including validity, reliability. Most importantly, Turkish EFL teachers were found to need training in assessing productive skills particularly speaking. In a more recent study, Yastibas and Takkac (2018) questioned the way language instructors improve LAL through interviews and discussions concluding that among three categories (previous assessment experience, assessment training and self-improvement) pre-service assessment

training had the most influence on their development of LAL. In terms of Turkish pre-service context, Kömür (2018) showed in his study that pre-service teachers ($n=49$) need to put their theoretical knowledge into practice as they self reported themselves as not competent enough to administer a sound assessment.

Through an extensive research, Şişman and Büyükkarcı (2019) reviewed 82 research studies and articles thoroughly with an attempt to get a big picture of the issue by categorizing the main purposes of studies such as LAL needs, LAL levels, LAL perceptions, assessment courses, teachers' beliefs, practices, and training needs in assessment, professional development in LAL, etc. among others. As they noted, most of the studies pay attention to language teachers' training needs in assessment as a result of insufficient training. As extensively investigated and reported, the pre-service language assessment training seems to be insufficient and lacks practical knowledge and opportunities for prospective teachers (Atay, 2017; Büyükkarcı, 2016; Hatipoğlu, 2015; Öz & Atay, 2017; Şahin, 2018). Most recently, on the contrary to the aforementioned studies above, Kaya and Mede (2021) explored LAL levels of 195 EFL instructors in their mixed methods study reporting a satisfactory level of knowledge of language assessment of participants while no evidence was found for the impact of certain factors such as experience and background.

In sum, though profoundly investigated, LAL studies vary on the basis of different conceptualization and description of the term (Fulcher, 2012), diverse perspectives of stakeholders (Malone, 2013), several epistemological paradigms (Inbar-Lourie 2008a) and assessment context (Inbar-Lourie, 2017). Therefore, Inbar-Lourie (2017) emphasized that as situational understanding of the local context and culture-specific knowledge make it essential to see language assessment from a wider window, “contextually relevant and diverse assessment practices” rather than testing-oriented LAL seem to be fostering. To this end, the following parts attempt to describe the relevant context of the current study by discussing LAL development through pre-service teacher training programs.

2.6. Assessment Literacy Training in Pre-service Teacher Education Programs

Assessing student language abilities is a challenging task for language teachers as it requires both theoretical knowledge and practical utilization (O'Loughlin, 2006). As it is highlighted by many leading figures in the literature, language teachers' limited understanding of assessment skills (Gardner & Rea-Dickins, 2001; Volante & Fazio, 2007) in addition to poor graded language tests (Alderson, 2005) reveal that there is an urgent need to investigate the pre-service training for language assessment literacy development in all possible angles. Hasselgreen, Carlsen, and Helness (2004) utilized a survey with an aim to reveal the assessment training needs of teachers in Europe; revealing that language teachers needed training in areas such as portfolio assessment, preparing classroom tests, peer and self-assessment, item writing, interviewing and rating among many other areas. In this sense, pre-service teacher education programs undoubtedly take the first place in terms of assessment literacy development. In the same vein, as DeLuca and Klinger emphasized an approach is needed which combines practice, theory and philosophy together in assessment education (2010).

More specifically, Volante and Fazio underlined in their study that a specific course in assessment is necessary to improve assessment knowledge (2007). The main difficulty lies in determining a suitable content (what to teach), the methodology to use (how to teach), and the underlying philosophy (why to teach). Therefore, "what, why and how questions" are the main elements to conceptualize a language assessment course in a pre-service teacher training program. In line with this thought, Davies (2008) conceptualises "skills, knowledge and principles" paradigm which shapes the assessment training in which skills suggest the core concepts such as item writing, statistics, test analysis; knowledge covers relevant background knowledge in assessment and measurement; and principles include proper use of language tests, fairness and ethics (Davies, 2008: 335).

Most recently, Fulcher (2020) proposed a model called "apprenticeship approach" to LAL development in which students of language testing are considered as "apprentices who learn by doing: using theory to design, research to create, values to assess" embedded in learning oriented approach (Fulcher, Panahi & Mohebbi, 2022: 52). Thus, "doing" as a social activity yields a collaborative environment in which learners (as apprentices) and

teachers (as mentors) cooperate, interact and reflect (Fulcher, 2020). In this sense, the apprenticeship classroom, as Fulcher calls it, is managed through a variety of tasks based on test design cycle (see Figure 1) focusing on different aspects of test development such as “item writing and review, prototyping, or the analysis of pilot data” (2020: 23) by necessitating responsibility for learners in practicing “theoretical considerations, practical research, and test building” (2020: 23).

In conclusion, in order to prepare teacher candidates for real classroom settings, bearing challenges in mind, LTA courses need to be analyzed and modified accordingly (Hatipoğlu, 2010). To fulfil this purpose, the next part is devoted to analyze language assessment courses in pre-service language teacher education programs in Türkiye.

2.6.1. Assessment Courses during Pre-Service Language Teacher Education in Türkiye

In Türkiye, the pre-service education of language teachers is carried out by Higher Education Council (HEC) which is responsible for managing all higher education institutions. This council determines the content of the educational programs, modifies the content when needed (Kırkgöz, 2007, 2009). The duration of teacher education is generally four year while some universities offer compulsory one-year English preparatory education. There are both compulsory and elective courses for students. The students generally have to take written exams, to name mid-term or final exam in addition to their projects, presentations, reflection reports. As for their practical classroom experiences, they have a practicum course which gives them the opportunity to observe real classroom settings, and further practice teaching. Upon graduation, teacher candidates take an exam which is called Kamu Personeli Seçme Sınavı (KPSS) to be able to work at state schools to teach different target group of students such as primary, secondary, and high school students.

With regards to assessment courses in language departments, certain prerequisite questions need to be posed which further draws the path to outline the content, objectives, teaching process and assessment of the course. Considering the content of a LTA course, determining the topics to include, the time to allocate to each topic, the selection of course

book bears crucial importance (Hatipoğlu, 2010, 2015, 2017). For this purpose, the opinions of all stakeholders including student teachers, teacher trainers, practicing teachers, test developers, policy makers are of vital importance so as to contemplate the basic patterns of language testing and assessment. Given the importance, language teacher education programs have been providing either compulsory or elective language testing courses comprising of a wide range of topics (Brown & Bailey, 2008) under various names such as “Language Testing and Assessment”, “Exam Preparation in Foreign Language”, “English Language Testing and Evaluation (ELTE, henceforth)”, to name a few.

In the faculty of educations in Türkiye, in general terms there is only one testing and assessment course (Ölçme ve Değerlendirme Dersi), in learners’ native language which is required to be taken generally at the third class spring term. However, in language teacher education faculties, in addition to the previous one, there is one more specific language assessment course which students generally take at their fourth year spring term (8th term). Emphasizing the importance of assessment, the course covers a wide range of topics including certain testing terms such as reliability, validity, practicality, different testing types (e.g.: proficiency, achievement, diagnostic, and placement tests, direct vs. indirect testing, discrete point vs. integrative testing, norm referenced vs. criterion referenced testing, objective testing vs. subjective testing, communicative language testing), alternative assessment methods such as written and oral exams, self-assessment, peer assessment, portfolio assessment, reflection papers, research projects, and so on (CoHE, 2018). The course also covers stages of test design and interpretation of test scores. However, it is undeniable that there is a huge gap between theoretical explanations and real life practices. Demonstrating this gap explicitly, Hatipoğlu (2010) carried out a study in order to evaluate “English Language Testing and Evaluation” (ELTE) course at a state university in Türkiye by asking the contribution of the topics covered within the course by employing questionnaires and interviews. The findings of her study revealed that “*the testing language skills/knowledge, reliability and validity*” were the most reported topics by the participants who claimed that these mentioned skills and topics would ensure them to “*prepare fair tests*”. Moreover, the deficiency of the course was found as the lack of enough practice opportunities by the participants. Based on the findings, Hatipoğlu (2010) draws the attention to the most salient concerns touching upon the role of pre-service

teacher training and previous experience of learners on their assessment beliefs and attitudes, inadequate knowledge about testing and language testing.

Having discussed the brief history of language teacher education in Türkiye, it is obvious that the ultimate goal of language assessment courses in pre-service teacher education programs is to equip prospective teachers with adequate knowledge about language testing and effective practicum opportunities. Within this purpose, the development of language assessment literacy of teachers is of vital importance, which is being questioned within the current dissertation. However, in real classroom settings language assessment may not be practiced as it is theorized. Therefore, this chapter will end by discussing language assessment practices in Turkish education system along with the recent developments and necessary skills for online language assessment.

2.6.2. Language Assessment Practices in Turkish Education System

This part serves as a brief history of language education system in Turkish context. The process of teaching and learning English as a foreign language has always been assumed a problematic area in Türkiye for ages (Uztosun, 2018). For this reason, there have been dramatic changes in language education policies (Kırkgöz, 2008), some of which have been stemmed from the local needs while some of them have been necessary to keep up with the language policy changes in the modern world. However, it is controversial whether all these changes or attempts would serve the purpose or not as classroom practices tell a different story. To be specific, as a consequence of 2005 policy change; constructivist learning approach and learner centeredness appeared along with process-oriented assessment (Topkaya & Küçük, 2010). As displayed in the reports of research studies with regards to the implementation of these changes; it can be noted that learning and teaching experiences in a real classroom setting did not yield expected results (Çankaya, 2015; Kırkgöz, 2009; Topkaya & Küçük, 2010). In 2012, a new education system, namely 4+4+4 was introduced by MONE, the focus of which was on speaking and listening skills which suggests the tendency towards constructivist and communicative approaches. Still yet, the teachers, active implementers of the program in classroom setting, did not promote students' speaking and listening skills (Gürsoy, Korkmaz, &

Damar, 2013), which causes a big gap between theory and practice. As a consequence of grammar based teaching, students fail to improve their communicative competence (British Council, 2013). Even though there has been an increasing tendency towards communicative language teaching in language classrooms, yet there is not still any progress about it, particularly within EFL context (Alam, 2017; Uztosun, 2018). Basically, the possible reasons for this failure were reported as teacher-centeredness and textbook orientation by Williams (2017).

Taken together, in-service teacher education in Türkiye would be assumed as debatable (Uysal, 2012; Uztosun, 2018) considering the ineffective classroom implementations. To illustrate, in accordance with the changes taken place in 2005, language teacher education programs also went through some changes in the following years (2005-2006). This new program emphasized Common European Framework of Reference for Languages (CEFR) as a starting point (Cosgun-Ogeyik, 2009). What is more, 4 + 4 + 4 policy put the student at the centre, thus supporting learner autonomy and reflective teaching. However, as reported by Uztosun (2018a), teachers' competences were found inadequate to employ the new 4 + 4 + 4 policy into practice, thus causing certain instructional problems including teaching techniques and assessment methods.

So far, the language education policy and the reflections of the policies over teacher education programs and language assessment practices have been detailed. Now, specifically an overview of recent online language assessment practices in addition to challenges will be covered in order to identify necessary knowledge, competences and skills needed for 21st century language classes, learners and teachers.

2.6.3. New Trends and Recent Language Assessment Challenges during Covid-19

This part of the study addresses online assessment with a specific focus on the challenges faced during the COVID-19 pandemic. In brief, this section is driven by the desire to discuss relevant assessment difficulties and coping strategies in various contexts.

In doing so, it would be helpful to comprehend the LAL from a new perspective by taking necessary language assessment skills into consideration to cope with recent challenges.

With the increasing popularity and predominance of technology, online education has gained greatest attention worldwide. However, the outbreak of a global pandemic, COVID-19, has caused several unplanned difficulties as a result of increasing number of deaths, closures of schools and businesses, and lack of social relationships (Erdoğan, 2020). Concerning the challenges, the most radical one has occurred in the field of education, thus instead of formal face to face education, new way of teaching during the pandemic has been characterised by online teaching or distance education (Cicillini & Giacosa, 2020; Karataş & Tuncer, 2020; König, Jäger-Biela, & Glutsch, 2020). In this sense, social distancing during Covid-19 has caused all educational activities from lecturing to assessment to take place in online platforms. Similarly, in Türkiye, since March 16, 2020 educational activities including lecturing, exams at all levels have been suspended (T.C. Sağlık Bakanlığı, 2020). As a result of this decision, the students and teachers have faced serious challenges during online education derived from technical problems or individual experiences. These exceptional experiences during Covid-19 reveal a great need for understanding the necessary skills, needs and new approaches for lecturing and assessing learners.

Even though the terminology with regards to online education varies, the basic premises can be outlined. In this regard, with an attempt to reach a full insight of the basic tenets of online teaching and learning process, Hampel and Stickler (2005) pinpointed seven competences, to name “ (a) *basic ICT competence*, (b) *specific technical competence*, (c) *awareness of constraints and possibilities*, (d) *online socialization*, (e) *facilitating communicative competence*, (f) *creativity, choice/selection*, (g) *development of own style*.” However, criticizing the limited literature about online teaching and learning, Compton (2009: 81) with a specific focus on language teaching presented a pedagogical framework based on 3 main premises labelling them as a) technology in online language teaching; b) pedagogy of online language teaching; and c) evaluation of online language teaching. Based on these facets of online education, it can be emphasized that teachers’

competences related to their ICT skills, online teaching pedagogy knowledge and self efficacy beliefs contribute to maximizing students' learning, thus increasing quality of online education. With regards to the benefits of online teaching and learning; flexibility (Smedley, 2010), self-directed learning (Elzainy, El Sadik, & Al Abdulmonem, 2020), time and place independence, suitability for larger groups of learners, possibility of redoing the activities more than once, enhancing higher order thinking abilities, instant feedback opportunities can be emphasized. To make use of these advantages, one needs to be competent enough in digital tools and technology use, thus having skills, knowledge and abilities to utilize technological tools and devices effectively and appropriately to optimize learning is significant. Therefore, a competent teacher or student is expected to use their skills to manage an online learning and teaching with ease by using problem solving skills, adjusting to challenges, working in collaboration, having ICT skills. However, as it has been experienced during the Covid-19 pandemic, ordinary assessment practices have turned into a total crisis as a result of technical, methodological and safety concerns, thus it needs a careful consideration. Accordingly, in Türkiye universities decided to conduct assessment practices through assignments and projects (YÖK, 2020) without questioning of the preparedness level of learners and teachers in terms of technical and pedagogical competences.

For these reasons, as noted by König, Jäger-Biela and Glutsch pre-service teacher education programs need to be designed to develop digital competence as using digital tools and having enough knowledge about new approaches and pedagogy for an effective online teaching is of great value to achieve a better and motivating online teaching (2020). Besides, evaluating the online learning process from different aspects bears great importance as new needs, practices and responsibilities have aroused during this pandemic.

In general terms, although there is a wide range of studies investigating online education from implementers' perspectives in the literature (Arslan, Bircan, & Eroğlu, 2019; Chen, Dobinson, & Kent, 2020; Çakır, Karademir, & Erdogdu, 2018), there is a gap focusing on online assessment challenges, coping strategies, new assessment techniques or questions suitable for online platforms during Covid-19. A recent study carried out by Taopan, Drajadi, and Sumardi (2020) pinpointed the challenges in EFL classrooms as information technology literacy, internet facilities and teachers' skills to manage engaging

tasks. Moreover, designed as a case study including data coming from survey and interview results, the study by Lie et al. (2020: 804) investigated online engagement, difficulties and practices concluding that the factors affecting online learning are “learners, teachers’ prior exposure to online learning, technological knowledge, pedagogical knowledge, and the support system”. As it is clear from this research ICT related courses in pre-service teacher education have a facilitating impact on prospective teachers’ positive attitudes and preparedness levels. Another study conducted by Cicillini and Giacosa (2020) addressed engagement and interaction concerns in online EMI classes during Covid-19 by employing both quantitative and qualitative analyses concluding two types of barriers as internal and external. Based on the findings, they reported that feelings of participants as internal barriers and technical issues (internet connection problems, disruptive noise) as external barriers have to do with the quality and successful online EMI class.

In Turkish context, on the other hand, attitudes of implementers, satisfaction levels, implementation process of online education during Covid-19 are the main focus areas of studies. To start, in an attempt to evaluate distance English courses in terms of online tools, materials, and implementation process, thus showing learners’ satisfaction levels, Erdoğan (2020) conducted a mixed methods research study using a questionnaire and structured interviews with 50 preparatory program students concluding that most of the students were content with online courses. By a different perspective, Karataş and Tuncer (2020) tried to explore the effect of online teaching on language skills at a pre-service teacher education program through a qualitative study gathering data from 118 pre-service EFL teachers providing evidence that online teaching was the most fruitful for writing ability and the least fruitful for speaking.

The brief literature presented above on online teaching and learning process during Covid-19 pandemic at the period of school closures shows that there is a need for investigating online assessment practices, methods and challenges as having a clear understanding of online assessment provides gainful insights about the principles and techniques in the design and implementation of online language assessment. Elzainy, El Sadik and Abdulmonem (2020) emphasized the perceived merits of online assessments as enhancing students’ autonomy and their critical thinking skills. In terms of feedback, it is

also advantageous as it provides timely feedback through an online test platform allowing students see their results afterwards. However, despite its promises, in most cases it is shown that online education does not provide teachers with mechanisms or platforms to assess learning meaningfully and give feedback to learners which further pose a stopping point in the continuity of learning and improvement (EAPRO, 2020). On the contrary to the benefits, online assessments have also drawbacks such as cheating and technical difficulties during exams. In order to cope with these difficulties, some solutions are offered by Elzainy, El Sadik and Abdulmonem (2020) as follows: time limit for each question and the whole test, designing brief exams, preparing scenario-based questions, randomising questions and answer choices, not permitting to go back to the former question among others.

Concerning online assessment integrity, cheating seems the prevailing concern. However it is noted that cheating during online assessment is not more common or easier than in traditional assessment (www.onlineeducation.com). Moreover, on the contrary to the extensive belief, online assessment methods can be monitored more easily through some web applications than traditional assessments (Krueger, 2015). But still, ensuring integrity in an online assessment bears great importance; therefore some actions need to be taken. So as to cope with the integrity concerns in online assessment supporting honesty comes to the fore. Furthermore, having a balance between formative assessment and summative assessment in addition to diverse question formats and types (multiple choice, open ended, true false, etc.) with a careful consideration of time limit ensure quality in online assessment, thus aiding in overcoming challenges to some extent.

2.7. Chapter Summary

Through historical and contextual background information, the rationale of the current research and theoretical framework were formulated to orient the necessary background of the study by discussing the key considerations in the development of LTA in Türkiye. Additionally, teacher beliefs about assessment, as a strong indicator of their classroom practices, have been detailed which suggests an overall picture about assessment practices in real classroom settings. Having outlined teacher beliefs extensively, it is

possible to iterate language assessment literacy and language assessment training from a critical perspective. Thus, having explained the domain of language assessment and assessment literacy in relation to pre-service language teacher education and assessment courses in Turkish context, the relevant literature review structured under three main headings such as language assessment beliefs; language assessment literacy and language assessment training were further detailed. By discussing prominent research, an overall picture about language assessment practices in real classroom settings was presented. Moreover, describing the main challenges of online language assessment during pandemic was underlined in the study with an attempt to demonstrate the new challenges and new needed skills about language assessment in today's world. Finally, having outlined language assessment literacy and its development extensively, it is possible to move to the methodology part of this study to shed light on the purpose, research design and the participants in Chapter Three.

CHAPTER THREE

METHODOLOGY

3.1. Chapter Introduction

Having presented the historical evolution of language testing and assessment through an extensive literature review mainly focusing on language assessment literacy and teacher beliefs in Chapter two, the study now goes on with Chapter three dealing with the research design and the purpose of the study. To this end, the introductory discussion starts with the purpose of the study with an emphasis on the research questions. Secondly, the research design and the rationale are highlighted by discussing the philosophy behind. Finally, data collection procedure and tools will be detailed by addressing ethical considerations, reliability and validity concerns aroused within the study.

3.2. The Purpose of the Study

The purpose of this study is to explore language assessment literacy levels and needs of participants and further to question their beliefs about language assessment. Moreover, language teacher education programs will be under scrutiny with regards to their contribution to student teachers' language assessment literacy development. To start, the LAL levels, needs and language assessment beliefs of in-service EFL teachers will be explored. Further, the study is planned to take place at a language teacher education program of a state university with the help of student teachers majoring at English language teaching department so as to explore the contribution of language assessment courses to their language assessment literacy levels and beliefs. Theoretically, the study intends to show the importance of LAL development in language teacher education as part of academic and professional success. The researcher wishes to make a contribution to the field of language testing and assessment by providing valuable information for various stakeholders including program developers, teachers, prospective teachers, and policy makers in designing an effective and practical LTA course to improve teacher candidates' language assessment literacy knowledge, skills and competences by covering necessary topics within this course.

3.3. Research Questions

To fulfil the purposes mentioned above, this study aims to answer the following research questions:

RQ 1: What are the language assessment literacy needs and levels of pre-service language teachers?

RQ 2: What are the language assessment literacy needs and levels of in-service language teachers?

RQ3: What are the pre-service and in-service language teachers' beliefs about language assessment?

RQ4: What is the contribution of language testing and assessment (LTA) course offered in pre-service language teacher education programs?

RQ5: What are the language assessment-related topics that participants recognise as important for their language assessment literacy development?

RQ6: What are the most challenging topics in terms of language assessment literacy?

RQ7: What are the possible ways for language assessment literacy development?

RQ8: What are the pre-service and in-service language teachers' perceptions about online language assessment?

3.4. Research Design and Rationale

A mixed methods research design based on convergent parallel design is employed in the study for analyzing the data coming from the volunteer participants' responses based on the questionnaires, interviews, and course documents. Creswell (2007, 2013) defined mixed methods as a process in which collecting, analyzing, and mixing quantitative and qualitative data are done in order to understand the research problem thoroughly. From a post positivistic perspective, the researcher attempted to investigate the issue from participants' own realities in their naturalistic contexts (Cohen, Manion, & Morrison, 2007) thus letting their voices be heard (Inbar-Lourie, 2017). Combining qualitative and quantitative data yield more comprehensive data thus increasing construct validity (Giraldo, 2021). To obtain the data, with an aim to explore the phenomenon under examination here, the researcher employs a survey in which a questionnaire (including Likert type questions), interviews (semi-structured, focus groups, follow up discussions), and related documents are utilized.

In this sense, qualitative data will be collected in the forms of interviews (individual and/or focus groups and e-mail interviews) while quantitative data will be reached through the questionnaire; thus achieving triangulation in terms of data type and analysis (Creswell, 2013). In doing so, this research design will be advantageous for its involving both qualitative and quantitative data for a comprehensive examination of the issue by complementing the missing sides inherent to each data set. Therefore, while the main purpose is to reach a deeper understanding and in-depth sights of the participants through interviews, the questionnaire aims at gaining an overall picture of the issue.

3.5. Research Site and Participants

As the main purpose of the study was to explore language assessment literacy of different stakeholders, purposive and convenience sampling was used. The target groups of participants consist of two main informants as pre-service EFL teachers and in-service EFL instructors working at state universities in Türkiye. More specifically, both quantitative and qualitative data were gathered from the pre-service EFL teacher candidates majoring in

their last year (4th year, 8th academic term) when language assessment related course is taught. As one of the main purposes of the study is to unveil the LAL needs and levels of pre-service EFL teacher candidates, they need to have taken the course before data collection. Thus, the researcher collected the data in May 2022 at the end of the spring term at a state university.

Upon reviewing the data, it was seen that one of the pre-service EFL teachers' data included missing values for the items, thus excluded from the data analysis and reports. Therefore, the detailed information of the 96 pre-service EFL teachers who took place in the study is displayed in Table 1 below.

Table 1

The information of the pre-service EFL teachers

		Number	Percent %
Gender	Female	62	63.9
	Male	34	35.1
Age	21	4	4.1
	22	27	27.8
	23	32	33.0
	24	8	8.2
	25-29	9	9.3
Perceived Level	Poor	2	2.1
	Fair	4	4.1
	Average	36	37.1
	Good	45	46.4
	Very good	9	9.3

As it is shown in Table 1, 96 pre-service EFL teacher candidates joined the study and filled out the questionnaire. The majority of them were female participants ($n=62$) while only one third of them were male participants (33.3 %). Their ages range from 21 years to 25 years and more while the highest percentage belongs to the group of participants with 23 years old (33.0 %). The majority of them ($n=45$) perceive their LAL level as good while only 6.2% of them consider themselves in either poor or fair LAL level.

Considering the second group of participants, Table 2 is prepared demonstrating the demographic information of in-service EFL teachers.

Table 2
Background information about in-service EFL teachers

		Number	Percent %
Gender	Female	72	73.5
	Male	26	26.5
Educational Degree	Bachelor	17	17.3
	Master in progress	3	3.1
	Master completed	22	22.4
	PhD in progress	46	46.9
	PhD completed	10	10.2
Teaching Experience	1-5 years	9	9.2
	6-10 years	19	19.4
	11-15 years	38	38.8
	16-20 years	15	15.3
	21- + years	17	17.3
Department of graduation	English Language Teaching	75	76.5
	English Language and Literature	12	12.2
	Linguistics	3	3.1
	Translating and Interpreting	6	6.1
	Other	2	2.0

In the second phase of data collection, in-service EFL teachers from universities either state or foundation were invited to take place in the study. Upon reviewing the data, it was seen that 4 in-service EFL teacher participants were not EFL instructors working at universities (they were high school or elementary school teachers), so their data were removed from the analysis. In the end, 98 in-service EFL instructors from 12 state and 5 foundation universities provided necessary data for the researcher. As they are the main implementers of assessment practices within classrooms, their responses are of vital importance. Their teaching experiences range from 1-5 years to 21 years and more while the highest percentage belongs to the group of participants with 11-15 years experience (38.8 %). The majority of them ($n=46$) have been studying for their PhD degree while 10

of them hold PhD. Concerning their department of graduation, it can be seen that the biggest percentage is 76.5 % for English Language Teaching department.

Moreover, the detailed information about the profiles of the participants who took place in interviews is demonstrated below with their abbreviations coded as “T” for EFL teacher participants and “S” for pre-service EFL teacher candidates.

Table 3

Interviewed in-service and pre-service EFL teachers’ profile

Code	Interview Mode	Length of Interview	Gender	Educational Degree	Teaching Experience	Department of Graduation
T1	Email	3 pages	Female	PhD on-going	6-10 years	ELT
T2	Email	4 pages	Female	PhD on-going	11-15 years	ELT
T3	Email	5 pages	Female	PhD on-going	11-15 years	ELT
T4	Email	4 pages	Female	PhD on-going	6-10 years	ELT
T5	Email	5 pages	Female	PhD on-going	11-15 years	ELT
T6	Email	3 pages	Female	PhD on-going	16-20 years	ELT
T7	Email	3 pages	Female	PhD	11-15 years	ELT
T8	Email	5 pages	Female	PhD on-going	16-20 years	ELT
T9	Online	38 min. 58 sec.	Female	PhD on-going	11-15 years	ELT
T10	Online	25 min. 58 sec.	Female	PhD on-going	1-5 years	ELT
T11	Face-to- face	45 min. 35 sec.	Female	PhD on-going	11-15 years	ELT
T12	Face-to- face	59 min.	Female	PhD	11-15 years	ELT
T13	Face-to- face	32 min. 16 sec.	Female	PhD	11-15 years	ELT
T14	Face-to- face	32 min. 16 sec.	Male	PhD on-going	11-15 years	ELT
S1	Online	28 min. 05 sec.	Male	Undergraduate	-	ELT
S2	Online	31 min. 40 sec.	Female	Undergraduate	-	ELT

As displayed above in Table 3, ninety-two percent of informants were female, and 61.5% of them fell in 11-15 years of teaching experience group; while 15.3% of female participants had 6-10 years of experience. In general sense, the majority of the interviewees (78.5%) have been studying for their PhD while the remainder (only 3) held a doctorate degree. The graduation department of the entire sample is ELT which is considered as an opportunity to get richer data as they are accepted familiar with language assessment knowledge and practices. As shown in the table, the interview mode had three options as face-to-face, e-mail and online. Only 14.2% of the sample joined in online interview session whereas the majority of them (57.1%) chose to take place in e-mail interview. The

length of interviews ranges either from 25 minutes to 1 hour or from 3 pages to 5 pages. Concerning with the interviews of pre-service EFL teachers, it was seen that 2 of them voluntarily contributed to qualitative data via Zoom which lasted approximately half an hour.

3.6. Data Collection Procedure

Employing a mixed methods research design, the data were collected from two sources (in-service EFL instructors and pre-service EFL teacher candidates) using three data collection tools (questionnaire, focus group discussions and/or semi-structured interviews, course documents), thus achieving data triangulation. However, as a result of pandemic, data collection became a challenging task, so e-mail interviews, video-conferencing tools and online platforms such as Google forms and WhatsApp or Zoom were used.

For qualitative data, semi-structured interviews, focus group discussions and follow-up interviews were done to capture a comprehensive picture of the issue as interviews provide in depth understanding (Cohen, et al. 2007), which undoubtedly enables the study with its strengths. Initially, the EFL instructors were contacted via e-mail or WhatsApp in December 2021 and January 2022 and asked to take place in the survey and focus group interviews, hence an interview invitation email or WhatsApp message to participants was sent. 14 in-service EFL instructors and 2 pre-service teacher candidates responded to this invitation and volunteered to participate the interview session. To frame the interviews, twenty two guiding questions were prepared (see Appendix 2) under three main titles such as “Assessment Beliefs”, “Needs for Assessment” and “Online Assessment”. Thus, the Google forms link of the questionnaire and interview questions for in-service language instructors were shared. E-mail, Zoom and WhatsApp interviews were employed to provide convenience for participants who were not able to schedule face to face interviews because of the pandemic. Therefore, the interviews were supervised as semi-structured and focus group and recorded in Zoom or voice recording program. The other qualitative data was obtained from language assessment course documents to analyze in terms of its contribution to EFL teacher candidates’ LAL levels for a deeper understanding by taking the content, objectives, outcomes, and assessment aspects of the

course into consideration.

Concerning the quantitative data, the questionnaire was designed in Google forms to collect responses from pre-service EFL student teachers and in-service EFL instructors separately. The online questionnaire in the present study was designed to find out the language assessment literacy levels and needs of participants and further to question the perceived importance levels of language assessment related topics taught and/ or expected to be taught at pre-service language teacher training programs, thus having three major parts. The Google forms link of “Language Assessment Literacy Questionnaire” was shared through e-mail, social media, WhatsApp and online lessons with a purpose to reach a wider population.

3.7. Development of the Instruments and Piloting Process

Concerning the design of the data collection instruments, the researcher analyzed the questionnaire and interview questions in terms of content (the wording of the questions, the order of the questions, meaning, etc.) and the format (number of the items, the length, subcategories, etc.). Moreover, some crucial steps were followed to pilot the data collection tools, in doing so, the researcher ensures clarity of the questions by overcoming misinterpretations or ambiguities of the items, thus attempting to reach valid and reliable data.

3.7.1. Development of the Questionnaire

Having reviewed the related literature profoundly, the questionnaire distributed to the participants in this study was adapted mainly from Fulcher’s survey (2012). The fourth question (Q4) in his study was used as a departure point in this thesis. It had 23 sub questions addressing language assessment related topics which were grouped under four categories on the basis of factor analysis results, to name a) Test design and development, b) Large-scale standardized testing, c) Classroom testing and washback, d) Validity and reliability. Additionally the Cronbach’s alpha values for each group were reported as follows: a) as .89, b) as .86, c) as .79, d) as .94. As a whole the reliability value was found as .93 indicating a high level of reliability. When it comes to the design of the questionnaire in this study, Likert-type questions involving abovementioned items in

Fulcher's study were grouped under three main parts identified as needs, levels, and perceived importance for 23 language assessment related topics. Therefore, Part 1, Part 2 and Part 3 included 5-point Likert-type questions to gather participants' opinions about the needs, levels and perceived importance of language assessment related topics. So, the first part seeks to answer the question about the levels of LAL including 23 sub-items while the second and third parts are related to needs and perceived importance for assessment related topics including the same items. Part 4, on the other hand is just about background information of the participants with a focus on LAL related experiences of participants such as previous training, etc.

In a detailed manner, for each question in **Part 1** the participants are asked to choose the options among "poor", "fair", "average", "good" and "excellent", while **Part 2** questions the need of participants in terms of LAL offering 5 options as "no need", "low", "moderate", "high", and "very high". The last part of the Likert type questions - **Part 3** - seeks to answer the perceived importance levels of assessment related topics to be covered in LTA course presenting 5 options as "Unimportant", "Not very important", "Fairly important", "Important", and "Essential". The questionnaire was shared on the Internet using Google forms link and as informants answered the questionnaire, their responses were automatically recorded on the web server, from which the researcher could download for use in SPSS.

3.7.2. Development of the Interview Questions

While constructing interview questions, an extensive literature review was carried out (Fulcher, 2012; Jeong, 2013; Tao, 2014; Tsagari & Vogt, 2017; Vogt & Tsagari, 2014) in order to identify the main themes, constructs, controversial points, gaps, perspectives, philosophies with regards to LAL and language assessment beliefs, thus designing a framework for the interview questions. Therefore, the questions mostly endeavoured to determine the features of a good language assessment, the reasons for assessment, and the most challenging topics of assessment, language assessment beliefs and the ways to improve language assessment literacy on the basis of participants' responses. Moreover, certain items were added with regards to online assessment as a global pandemic (Covid-

19) has had a lasting effect on language assessment all over the world. These additions included items related to conducting online language assessment, online feedback, online assessment methods and techniques along with reliability and validity considerations in online assessment. The questions are designed as semi structured with an aim to modify them if necessary to gain more insights and rich data (Dörnyei, 2007; McDonough & McDonough, 2006). In synthesis, the interview questions are expected to unveil language assessment beliefs and needs of the participants along with their opinions related to online assessment. The interviews were done either face-to-face or online due to the pandemic and recorded lasting for about 40-60 minutes.

Having constructed the first version of the interview questions, to validate it, 3 language instructors holding PhD degrees from 3 different state universities were consulted and asked to share their opinions on each question and format of the interview from a critical eye as an insider (EFL teachers) and outsider (as if being a participant) at the same time. Based on their comments and critics, certain amendments and modifications were done to ensure clarity and meaning for the respondents, thus enhancing the readability, clarity, and meaningfulness of the questions by ruling out misunderstandings and ambiguity (Dörnyei, 2003). Secondly, to ensure the internal reliability and reduce the misunderstandings and ambiguity within the questions, the interview procedure was piloted with one colleague who was chosen by convenience sampling. The same procedure was followed, namely the interview was done face to face and lasted for 35 minutes. However, for actual participants some semi-structured interviews were planned to be conducted as e-mail interviews or focus group interviews on Google Meet or Zoom.

All in all, the questions were modified in terms of both its content and layout (headings, sub questions, etc.) on the basis of suggestions offered by aforementioned experts. Some modifications were related to the length of the questions or word choices while some other suggestions were done related with the ordering of the questions. For example, two questions under the beliefs of assessment category were changed as follows: “How do you think assessment affects student learning?” and “How do you think assessment affects your teaching?” as the wording of these questions was found

misleading. That's to say, the word "improves" in the first version was changed as "affects" as it was believed to lead the interviewees. Besides, to avoid ambiguity, "what are the features of a sound assessment" was changed as "what are the features of a good-sound assessment?" Concerning second part of the interview with regards to "Needs for Assessment", two questions were modified to get richer data by adding "If yes, How? If not, why not?", thus follow up explanations could be obtained. What's more, the order of the questions was also changed by taking one of the experts' views into account. She suggested that it would be a better idea to start with the reasons of assessment. In conclusion, a total of 22 open ended questions under 3 categories were constructed as the framework for the interviews.

3.7.3. Course Documents

With an attempt to gain more insights of LAL development, language assessment course syllabus of the state university from which the data was collected was analyzed in terms of its content, objectives, teaching and learning process, outcomes and assessment dimensions. Course syllabus was reached from the university website while details were discussed with participants. Although the name and the content of the specific LTA course may vary among universities, in order to achieve data triangulation the content, the materials used during teaching and learning process, the stated objectives, the outcomes, and assessment of the course within the syllabus were examined in detail through document analysis.

3.8. Data Analysis

3.8.1. Qualitative Data

Deciding on analytic angle for qualitative data analysis; content analysis was found suitable for drawing conclusions from the gathered data. Thus, the qualitative data obtained from interview recordings were transcribed and went through content analysis by utilizing open coding based on the guidelines by Creswell (2013), a protocol proposed by Bryman's (2015) as well as taking trustworthiness criteria into account (Lincoln & Guba, 1985). In this sense, the content analysis of qualitative data covered the following steps:

The first step requires preparing and organizing the gathered data for analysis. Therefore, the recorded interviews either audio taped or recorded on Zoom were transcribed by the researcher. Secondly, transcribed data were reviewed by reading several times and explored any recurring phrases or similar ideas, thoughts, concepts, phrases, etc (Strauss, 1987). Then, for categorizing similar ideas and key terms, open coding was applied by demonstrating the connected idea. Open coding is employed using an inductive approach by grouping codes determining the most common themes (Charmaz, 2014; Weber, 1990). As a fourth step, reviewing the codes, the coded data was revised and written as themes. The common emergent themes and topics were identified and grouped in major categories by sorting in line with the interview questions (Creswell & Poth, 2018; Dornyei, 2007), thus the emergent data align well with the guiding questions within the interviews. Lastly, the conclusions drawn from the data set were presented cohesively based on the recurring themes and opinions (Creswell, 2007, 2013) through tables and charts. Both employing inductive approach and deductive approach (Creswell, 2009), the researcher kept some codes in her mind based on interview questions while some other codes and themes emerged from the respondents' opinions and discussions. Concerning interpretations of the data verbatim quotations were also included to support the emerged themes and codes.

In order to achieve inter-coder reliability, thus ensuring more meaningful interpretation of the data, initially a subset of data belonging to each group of participants was chosen so that the data should represent entire data set. The same data set was reviewed and coded by a different coder in order to enhance inter-coder reliability of qualitative data. To this end, the formula by Miles and Huberman (1994) was used in the following way: $\Delta = C \div (C + \partial) \times 100$. Initially, Δ stands for the reliability value; C represents the number of agreements by two coders while ∂ stands for the number of disagreements. The expected reliability value for inter-coder reliability is considered as 80% and more. Within this scope, the inter-coder reliability of the qualitative data in this current study was found as 84.9% which was calculated in the abovementioned formula by Miles and Huberman (1994) as $\Delta: 135 \div (135 + 24) \times 100$ suggesting a satisfactory inter-coder agreement. The misalignment was tried to be solved through discussion and when a consensus was achieved on the interpretations of the responses, the final theme was listed.

Once the initial coding was done, relevant patterns were identified. In this regard, the generated themes drawn from the data set addressed the defining qualities of a good assessment, the reasons for assessment, and beliefs about language assessment, benefits and drawbacks of online language assessment, the perceived level of language assessment literacy of respondents. More importantly, in pursuit of ensuring reliability and validity of the qualitative data, the qualitative findings were reported both in numbers (frequencies and percentages) and quotations by following Miles and Huberman's point of views in their own words: "to keep yourself analytically honest, protecting against bias" (1994: 253).

What's more, the documents of LTA course (Test Preparation in English Language Teaching in this case) were analysed through content analysis to identify the basic tenets such as content, objectives and the outcomes of the course. Finally, findings of the quantitative data gathered via LAL questionnaire were discussed together with responses obtained from interviews and course documents to report on a wealth of insights with regards to the phenomenon under examination.

3.8.2. Quantitative Data

Quantitative data obtained through the LAL questionnaire was managed and analyzed using SPSS software (version 20) with calculations of descriptive, frequencies and percentages to determine the LAL levels and needs of informants along with the importance levels of LAL topics. Before carrying out calculations, the data was checked whether the data was normally distributed or not, thus proved it was reasonable to perform the statistical tests. To this end, normality test was computed and the Shapiro-Wilk test result was found $p=.445$ while Kolmogorov-Smirnov test result with a $p=.200$. Based on these outcomes, it was acceptable to compute parametric tests.

3.9. Validity and Reliability Concerns

The present study is designed as a mixed-methods study (Creswell, 2013) to shed light on the language assessment literacy levels, needs and beliefs of different stakeholders, thus achieving triangulation in terms of data source, data type and analysis. Enhancing

validity and reliability is aimed by triangulating data collection through questionnaire, interviews and document analysis.

In this sense, addressing reliability and validity measures in this current study is crucial as the researcher attempts to overcome the potential threats such as researcher bias, mono-operation, mono-method bias, inadequate theory. As Maxwell (2010) pinpointed, some methods to enhance validity including “obtaining rich data, validating respondents’ responses, data triangulation, peer debriefing, and member checking” were used within the study. In the same vein, thick descriptions of research context and participants, prolonged engagement with data are among the other measures to address the trustworthiness of the qualitative data (Lincoln & Guba, 1985; Mackey & Gass, 2005; Shenton, 2004).

With regards to validity, concerning data collection tools the researcher tried to achieve construct validity by consulting colleagues about the items of the interviews. To eliminate unreliable measurement threat to validity as put forward by Cook and Campbell (1979, cited in Lynch, 1996), a piloting process was followed for the validity of the instrument. By providing cross-references to other relevant research studies in the literature review and adequate theory, inadequate theory threat was overcome. Moreover, by triangulating the data collection tools (questionnaire, interviews and document analysis), the researcher also minimized mono-operation and mono-method bias threats to external validity.

When it comes to reliability, the interview transcripts were reviewed and coded by two different coders in order to enhance inter-coder reliability of qualitative data while the reliability of the questionnaire was reported as .93 (Cronbach Alpha value) by Fulcher (2012) suggesting a high internal consistency reliability. On the other hand, for external validity, as the number of the participants is limited to only 96 EFL student teachers and 98 EFL teachers, the generalizability of the findings is questionable. In the same vein, by providing an in-depth theoretical basis for the background to the language assessment literacy within the extensive literature review, the researcher handled the threats of inadequate theory to external validity. In terms of sampling procedure in the study, the researcher used convenience sampling which minimizes the threats regarding with the participants such as history, maturation, mortality to some extent.

Taking trustworthiness criteria into consideration (Lincoln & Guba, 1985), this current study aimed at achieving reliability and validity of data analysis ensuring 4 criteria as credibility, transferability, dependability, confirmability.

3.10. Ethical Considerations

To carry out a research study, qualitative or quantitative, the researcher is expected to follow ethical guidelines from the very beginning of her / his research to reporting results.

Therefore, the researcher sought to receive approval to ensure ethical concerns within this current study before collecting data; therefore, she submitted necessary documents of the research project including informed consent, ethics examination form, the form of questionnaire and interview questions to the Ethics Committee of COMU School of Graduate Studies. The ethics committee checked the research project documents consisting of the main research materials and data collection instruments whether they are ethically acceptable. As the researcher got permission from the institute, she started to collect data from April 2022 to June 2022.

With an attempt to handle researcher or participant bias and other potential ethical problems in the study, the study was designed by following some ethical guidelines. Firstly, as the qualitative part of the current study tries to elicit meanings and answers of the respondents' feelings and opinions, the researcher followed the principle of "voluntary participation" and "no harm" to participants. To ensure confidentiality, the researcher used abbreviations for the names of people, thus achieving anonymity. Finally, in order to avoid misinterpretations of the responses of participants, open and honest communication was employed. Therefore, the researcher aimed to overcome biases and misunderstandings. Otherwise, the credibility and reliability of the current paper would become problematic and questionable.

3.11. Chapter Summary

Having stated the purpose and research questions; the design of the study was described in detail presenting the development of data collection tools and piloting phase, the participants, data collection and analysis procedures within Chapter Three. Additionally, ethical concerns, reliability and validity checks were discussed to prove the trustworthiness of the study. Now, the next chapter, Chapter Four, will serve as a lens to the issue investigated in the study by reporting the findings of qualitative and quantitative data.



CHAPTER FOUR

RESULTS

4.1. Chapter Introduction

Functioning as a mirror, this chapter is aimed at probing the results of qualitative and quantitative data. The results are detailed in accordance with the relevant research questions orienting a clear understanding of the issue under investigation here. Thus, it starts with research questions and goes on with the main findings. Displaying the results of the study will be accompanied by an in depth discussion under each research question to draw attention to the salient points. By doing so, comparing the findings of the study with the relevant literature tries to portray the issue in a detailed manner. Thus, the chapter includes an extensive discussion mentioning a wide range of research studies.

4.2. Results and Discussion for Research Question 1

What are the language assessment literacy needs and levels of pre-service language teachers?

Even though the majority of the studies have a particular focus on various stakeholders including teachers, teacher trainers, language testers, language researchers, etc. in terms of their language assessment literacy (LAL, henceforth) levels and needs, still pre-service teacher candidates should be paid more attention. Remarkably, their LAL development is of vital significance to be investigated thoroughly as the main duty of pre-service teacher training programs is to equip prospective teachers with necessary skills, knowledge, and competences with respect to language assessment. For this reason, to fill the gap and provide an insightful perspective for the LAL development, this study attempts to hear pre-service EFL teacher candidates' voices, which in turn enables to find possible ways and solutions for LAL development of them.

While the gathered quantitative data is displayed through tables to reach a clear understanding of the issue from pre-service EFL teachers' point of views discussing the levels and needs in terms of language assessment related topics; the qualitative data obtained through semi structured interviews (one-to-one) provides detailed and in-depth interpretations about their language assessment beliefs, the possible reasons for a good or poor assessment based on their self responses and comments which will be cited in quotations.

Initially, in pursuit of exploring LAL levels of pre-service EFL teachers, some statistical measures including descriptive and frequencies were calculated through SPSS (20 Version) and the results are displayed below in Table 4 by illustrating the related items.

Table 4

Pre-service EFL teachers' perceived LAL levels

LAL Items	Poor		Fair		Average		Good		Excellent	
	N	%	N	%	N	%	N	%	N	%
History of Language Testing	32	33.0	25	25.8	30	30.9	9	9.3	1	1.0
Procedures in language test design	2	2.1	12	12.4	25	25.8	50	51.5	8	8.2
Deciding what to test			1	1.0	12	12.4	47	48.5	37	38.1
Writing test specifications/blueprints	2	2.1	11	11.3	29	29.9	36	37.1	19	19.6
Writing test tasks and items	1	1.0	5	5.2	27	27.8	41	42.3	23	23.7
Evaluating language tests			4	4.1	16	16.5	50	51.5	27	27.8
Interpreting scores			3	3.1	24	24.7	47	48.5	23	23.7
Test analysis	2	2.1	7	7.2	34	35.1	41	42.3	13	13.4
Selecting tests for your own use			3	3.1	19	19.6	50	51.5	25	25.8
Reliability	1	1.0	3	3.1	20	20.6	49	50.5	24	24.7
Validation	1	1.0	2	2.1	24	24.7	44	45.4	26	26.8
Use of statistics	8	8.2	27	27.8	26	26.8	19	19.6	17	17.5
Rating performance tests (speaking/writing)	1	1.0	10	10.3	27	27.8	51	52.6	8	8.2
Scoring closed-response items	1	1.0	8	8.2	25	25.8	41	42.3	22	22.7
Classroom assessment	2	2.1	3	3.1	25	25.8	39	40.2	28	28.9
Large-scale testing	4	4.1	21	21.6	36	37.1	28	28.9	8	8.2
Standard setting	2	2.1	13	13.4	31	32.0	37	38.1	14	14.4
Preparing learners to take tests			8	8.2	15	15.5	44	45.4	29	29.9
Washback on the classroom			4	4.1	23	23.7	39	40.2	31	32.0
Test administration	2	2.1	8	8.2	19	19.6	39	40.2	28	28.9
Ethical considerations in testing			10	10.3	20	20.6	31	32.0	36	37.1
The uses of tests in society	2	2.1	13	13.4	30	30.9	33	34.0	19	19.6
Principles of educational measurement			16	16.5	31	32.0	29	29.9	21	21.6

In response to the question about LAL levels, the quantitative data above demonstrates that the percentages of the EFL teacher candidates expressing excellent level with respect to the items “Deciding what to test “ and “Ethical considerations in testing” were the highest with the percentages 38.1 % and 37.1 % respectively. On the contrary, the lowest percentage belongs to the “History of Language Testing” with 1.0 % which shows

an alignment with their level for the same item under “Poor” category. That’s to say, 33.0 % of the pre- service EFL teachers ($n=32$) self reported that they have a very poor level for the historical evolution of language testing. Moreover, when “Poor” and “Fair” categories closely analyzed, it might be seen that the mostly mentioned items were “History of Language Testing” and “Use of statistics”, 58.8 % and 36 % respectively.

The data above indicates that the percentages of the EFL teacher candidates expressing good level with respect to the items “Rating performance tests (speaking, writing)” and “Procedures in language test design” and “Deciding what to test” were very high 52.6 %, 51.5 % and 48.5 % respectively. The mostly mentioned items with related to average levels, on the other hand, can be seen as “large scale testing” (37.1%) and “test analysis” (35.1%) which are in accordance with their training needs. These items are followed with “standard setting” and “history of language testing” with the percentages of 32.0 and 30.9 respectively. In general sense, the lowest mean value was found as 2.20 for “History of Language Testing” while the highest value was found as 4.24 for “Deciding what to test” based on the responses of pre-service EFL teachers. Similarly, a great number of participants reported a very high need of training for “History of Language Testing” (35.1%) and “Use of statistics” (28.9%) which they self reported to have low level. Likewise, findings emerged from qualitative data obtained through interviews with pre-service EFL teachers proves an alignment between the items which were labelled in a very high need for training category and low level of LAL in the questionnaire.

To put it short, almost half of the participants (46.4%) mentioned they have good level with respect to LAL while 37.1 % reporting average level of LAL. Similarly, Kaya and Mede’s study (2021) reported a satisfactory level of LAL in their mixed method study with 195 participants. Yet, the findings above contradict with the results as put forward by many leading figures in the literature, concluding low assessment literacy levels for teachers (Hasselgreen, Carlsen & Helness, 2004; Tsagari & Vogt, 2017; Volante & Fazio, 2007; Xu & Brown, 2016), thus language assessment courses offered at teacher education programs are criticized as being too theoretical rather practical. In similar vein, many prominent research studies underline that assessment courses offered at pre-service teacher

education program do not improve learners practical skills as they are mostly reported being theory-oriented (Kaya & Mede, 2021; Ölmezer-Öztürk & Aydın, 2018). More comprehensively, Ölmezer-Öztürk and Aydın (2018) concluded that EFL teachers have limited and low LAL levels in terms of assessing productive skills while they self reported themselves as having high LAL level in assessing receptive skills such as reading, grammar, etc. However, within this study, the pre-service EFL teachers self reported that they have a good level for “Rating performance tests (speaking, writing)” with 52.6 %. As an eye catching result, this can be explained by the positive impact of LTA course the pre-service EFL teachers have just taken which is also supported by DeLuca and Klinger’s study (2010) with 288 teacher candidates concluding that assessment course had considerably positive impact on their assessment literacy development.

When it comes to LAL needs of pre-service EFL teacher candidates, it is reasonable to interpret the data bearing the LAL levels in mind as it is anticipated that if the LAL level is good or excellent, then the LAL need should be in low or no need category. In order to discuss the needs of the first group of participants, Table 5 is prepared below.

Table 5

Pre-service EFL teachers' perceived LAL needs

LAL Items	No		Low		Moderate		High		Very high	
	N	%	N	%	N	%	N	%	N	%
History of Language Testing	4	4.1	15	15.5	27	27.8	17	17.5	34	35.1
Procedures in language test design	8	8.2	30	30.9	22	22.7	22	22.7	15	15.5
Deciding what to test	26	26.8	21	21.6	21	21.6	13	13.4	16	16.5
Writing test specifications/blueprints	12	12.4	23	23.7	14	14.4	21	21.6	27	27.8
Writing test tasks and items	15	15.5	25	25.8	24	24.7	17	17.5	16	16.5
Evaluating language tests	20	20.6	26	26.8	21	21.6	17	17.5	13	13.4
Interpreting scores	12	12.4	23	23.7	24	24.7	27	27.8	11	11.3
Test analysis	11	11.3	24	24.7	23	23.7	28	28.9	11	11.3
Selecting tests for your own use	20	20.6	30	30.9	21	21.6	15	15.5	11	11.3
Reliability	22	22.7	23	23.7	17	17.5	14	14.4	21	21.6
Validation	21	21.6	19	19.6	21	21.6	14	14.4	22	22.7
Use of statistics	15	15.5	10	10.3	18	18.6	26	26.8	28	28.9
Rating performance tests (speaking/writing)	12	12.4	21	21.6	21	21.6	28	28.9	15	15.5
Scoring closed-response items	16	16.5	28	28.9	18	18.6	20	20.6	15	15.5
Classroom assessment	17	17.5	21	21.6	19	19.6	27	27.8	13	13.4
Large-scale testing	10	10.3	16	16.5	25	25.8	30	30.9	16	16.5
Standard setting	16	16.5	18	18.6	25	25.8	25	25.8	13	13.4
Preparing learners to take tests	22	22.7	23	23.7	20	20.6	16	16.5	16	16.5
Washback on the classroom	27	27.8	20	20.6	17	17.5	21	21.6	12	12.4
Test administration	20	20.6	24	24.7	21	21.6	21	21.6	11	11.3
Ethical considerations in testing	25	25.8	23	23.7	15	15.5	23	23.7	11	11.3
The uses of tests in society	17	17.5	26	26.8	25	25.8	16	16.5	13	13.4
Principles of educational measurement	14	14.4	19	19.6	18	18.6	28	28.9	18	18.6

As the results indicate, the majority of the pre-service EFL teachers self reported high or a very high need for the following LAL items as “Use of statistics” (55.7%) and “History of Language Testing” (52.6%). The highest percentages of “no need” for training in LAL topics belonged to the items of “Deciding what to test” (26.8%) and “Washback on classroom” (27.8%). When analyzed closely, these results portray that the pre-service EFL

teachers feel themselves ready to decide what to test and manage the washback effect of language assessment while most of them anticipate the difficulty of using statistics in language assessment. Moreover, the results showed high need in knowledge regarding history of language testing, but relatively lower need in ethical concerns in testing which are also in alignment with their perceived self stated LAL levels. The qualitative data emerged from the content analysis of course documents revealed that “washback-backwash” and “security and transparency” were considered as the corner stones of language assessment which in turn enhanced prospective teachers’ competencies about the impact of assessment and ethical considerations in a language test which were also supported by the quantitative data above.

Qualitative data on the other hand was used to reveal the specific needs in terms of knowledge, skills and competences based on the two interview questions, to name “Do you feel competent when assessing students? If yes, how? If not, why not?” and “What do you think about your needs (knowledge, skills, practice) in assessment process?” Considering the participants’ opinions quoted below, the quantitative data and qualitative data seem in an alignment, thus supporting each other. As the quantitative data displayed 70.1% of the pre-service EFL teachers regarded their LAL needs as moderate and/or high category which corroborates the following quotations. For instance, S1 acknowledged that:

I need to improve myself in this sense. Because I didn’t have an opportunity to practice language assessment in practicum course, I couldn’t benefit enough from the course (S1).

Moreover the following comment can be noted as a characteristic of a considerable number of participants within the present study:

I am not qualified enough due to the overexposure to multiple choice tests, as a result I do not have necessary knowledge about alternative assessment types, tasks (S2).

Of all the related LAL items within the questionnaire, "Use of statistics" has the highest percentage in terms of participants' training needs. Namely, 55.7 % of the pre-service EFL teachers responded that they were in a very high need for "use of statistics" during language testing and assessment procedure. On the other hand, 15.5% noted they were in no need for such training as they felt competent in statistics. Moreover, nearly half of the participants ($n=43$) reported high or very high need for training in order to be able to rate performance tests (speaking/ writing). When compared to other studies within the literature, the results above can be said to be in line with the findings of Ölmezer-Öztürk and Aydın's study (2018) where it was concluded that EFL teachers have limited and low LAL levels in terms of assessing productive skills while they self reported themselves as having high LAL level in assessing receptive skills such as reading, grammar, etc. With respect to preparedness level of pre-service teachers, Kömür (2018) carried out a study concluding that the participants ($n=49$) need to put their theoretical knowledge into practice as they self reported themselves as not competent enough to administer a sound assessment. More recently, Kavaklı and Arslan (2019) stated pre-service language teachers need authentic courses and opportunities for real classroom practices, which would allow them to put theoretical knowledge into practice.

In broad terms, considering the overall picture of LAL needs of pre-service EFL teacher candidates, it might be portrayed that the majority of student participants ($n=73$) are in moderate, high or a very high need (74.3%) on the basis of their self responses. When the results are analyzed closely, the lowest need for training was concerned with the item "Selecting test for your own use" with the mean value of $M=2.66$ ($SD=1.282$) while the highest value for training was $M=3.64$ ($SD=1.226$) for "History of language testing". On the contrary to their LAL levels they have self reported as good, still yet nearly all pre-service EFL teachers regarded themselves in need for further LAL training. Considering qualitative data, the most eye catching and commonsense opinion about LAL need is noted by one of the participants as the need of knowledge for employing alternative assessment and productive based tasks (S1).

As Şişman and Büyükkarcı (2019) noted, most of the studies pay attention to language teachers' training needs in assessment as a result of insufficient training. As extensively investigated and reported, the pre-service language assessment training seems to be insufficient and lack practical knowledge and opportunities for prospective teachers

(Atay, 2017; Büyükkarcı, 2016; Hatipoğlu, 2015; Öz & Atay, 2017; Şahin, 2018). However, it should not be forgotten that, on the contrary to what theory based teaching offers, practical language assessment training provides opportunities for hands-on experiences to bridge theory-practice gap as active implementers of assessment activities requires to know how to put into practice the assessment related theories.

4.3. Results and Discussion for Research Question 2

What are the language assessment literacy needs and levels of in-service language teachers?

For further consideration, the second research question attempts to map the LAL of in-service EFL teachers by stating their LAL needs and levels. Although different profiles and stakeholders have been investigated in terms of their needs, levels and perceptions about LAL, but still teachers take the first place having a central role as they are the main actors in doing language assessment (Giraldo, 2018; Tsagari & Vogt, 2017). To put it simply, their LAL development conveys importance as this means enhancing students learning and improving the teaching instruction at the same time.

On the one hand, the quantitative data displays the levels and needs of in-service EFL teachers in terms of language assessment related topics through tables to reach a clear understanding of the issue; the qualitative data gathered through semi structured and focus groups interviews, on the other hand, enables a valuable bulk of ideas and thoughts about their language assessment beliefs, the possible reasons for a good or poor assessment, the ways for LAL development, the contribution of language assessment course at their pre-service teacher training based on their self responses and comments.

In an attempt to find out the current LAL needs, some statistical calculations were employed through SPSS (20 Version) concluding very high need for some LAL topics in

the questionnaire as demonstrated in Table 6 below which is similar to pre-service EFL participants within the present study.

Table 6

In-service EFL teachers' perceived LAL needs

LAL Items	No		Low		Moderate		High		Very high	
	N	%	N	%	N	%	N	%	N	%
History of Language Testing	14	14.3	28	28.6	19	19.4	24	24.5	13	13.3
Procedures in language test design	4	4.1	17	17.3	21	21.4	26	26.5	30	30.6
Deciding what to test	12	12.2	22	22.4	21	21.4	13	13.3	30	30.6
Writing test specifications/blueprints	4	4.1	19	19.4	21	21.4	27	27.6	27	27.6
Writing test tasks and items	8	8.2	16	16.3	21	21.4	24	24.5	29	29.6
Evaluating language tests	13	13.3	16	16.3	17	17.3	23	23.5	29	29.6
Interpreting scores	9	9.2	15	15.3	20	20.4	21	21.4	33	33.7
Test analysis	5	5.1	15	15.3	22	22.4	25	25.5	31	31.6
Selecting tests for your own use	15	15.3	24	24.5	14	14.3	21	21.4	24	24.5
Reliability	10	10.2	19	19.4	14	14.3	27	27.6	28	28.6
Validation	11	11.2	16	16.3	15	15.3	27	27.6	29	29.6
Use of statistics	6	6.1	20	20.4	18	18.4	28	28.6	26	26.5
Rating performance tests (speaking/writing)	11	11.2	14	14.3	22	22.4	21	21.4	30	30.6
Scoring closed-response items	18	18.4	19	19.4	21	21.4	17	17.3	23	23.5
Classroom assessment	8	8.2	23	23.5	21	21.4	19	19.4	27	27.6
Large-scale testing	6	6.1	17	17.3	23	23.5	22	22.4	30	30.6
Standard setting	10	10.2	13	13.3	25	25.5	24	24.5	26	26.5
Preparing learners to take tests	13	13.3	19	19.4	21	21.4	21	21.4	24	24.5
Washback on the classroom	11	11.2	25	25.5	17	17.3	23	23.5	22	22.4
Test administration	19	19.4	22	22.4	16	16.3	20	20.4	21	21.4
Ethical considerations in testing	14	14.3	26	26.5	14	14.3	21	21.4	23	23.5
The uses of tests in society	7	7.1	22	22.4	24	24.5	26	26.5	19	19.4
Principles of educational measurement	9	9.2	22	22.4	20	20.4	22	22.4	25	25.5

As shown in Table 6 above, in response to the question about LAL needs, almost one third of the informants stated that they have a very high need for training for the following

LAL topics: “Interpreting scores” (33.7%), “Rating performance tests (speaking/writing)” (30.6%), and “Large scale testing” (30.6%). As for the items with respect to the main issues in testing, “reliability” and “validity”, 27.6 % of the in-service EFL participants perceived their needs as high in addition to almost 30% of them considered their need in “very high” category. Additionally, as the results demonstrate the highest percentages of no need for training in LAL topics belonged to the following items: “Test administration” (19.4%) and “Scoring closed-response items” (18.4%). In general sense, the highest mean values for the need of training belong to the “Procedures in test design” ($M=3.62$, $SD: 1.206$) and “Test analysis” ($M=3.63$, $SD: 1.222$), while the lowest one is related to the “Scoring closed-response items” ($M=3.08$, $SD: 1.434$).

When compared to other studies within the literature, the results above can be said to be in line with the findings of several studies (Chung & Nam, 2018; Hasselgreen, Carlsen & Helness, 2004; Kremmel & Harding, 2020; Ölmezer-Öztürk & Aydın, 2018; Vogt & Tsagari, 2014). To start, Hasselgreen, Carlsen and Helness (2004) examined the assessment training needs of teachers in Europe in their study revealing that language teachers needed training in areas such as portfolio assessment, preparing classroom tests, peer and self-assessment, item writing, interviewing and rating among many other areas, thus it was noted that the training did not fully prepare participants to perform assessment-related activities. These findings also agree with Ballıdağ and İnan-Karagül’s study (2021) which identified the needs of EFL teachers as using self-or peer-assessment testing and assessing aspects of culture, using statistics. The least need for further training was observed in testing and assessing grammar and vocabulary in the same study which was also supported by the findings of this current study. Likewise, Ölmezer-Öztürk and Aydın (2018) concluded that EFL teachers have limited and low LAL levels in terms of assessing productive skills while they self reported themselves as having high LAL level in assessing receptive skills such as reading, grammar, etc. In line with this, the current study concluded that the in-service EFL teachers self reported that they are in a very high need of training for “Rating productive skills (speaking/writing)”.

Qualitative data on the other hand was used to reveal the specific needs of in-

service EFL teachers which attempted to identify as “knowledge, skills and competences” based on the following interview questions as “Do you feel competent when assessing students? If yes, how? If not, why not?” and “What do you think about your needs (knowledge, skills, practice) in assessment process?” When qualitative data analyzed, considering the majority of responses, it can be noted that interviewees generally found themselves competent (10 out of 14) in terms of language testing and assessment thanks to their educational background. To exemplify;

Yes I do, I got testing classes at the university and I went on Masters Degree, Phd, and got seminars about testing and assessment (T1).

Yes, because thanks to my education, I know, when, how, and what to assess, and what to do with these results (T3).

As it is clear from the utterances above, the interviewees think that they owe their being competent to their educational background at pre-service training or post graduate studies. One other participant also indicated the contribution of experience at testing office unit at a state university as in her own words:

I had training about how to write test specifications and about item writing and worked in evaluation and assessment unit of my organization, *thus* I have had theoretical and practical background knowledge (T5).

Concerned with the negative responses for the same question, the following comment is very crucial as it sheds light on the basic underlying problems with language testing and assessment training during pre-service education:

A clear answer, no. Because the LTA course was presented in the last term (8th academic term, spring time) when I had to take KPSS, ALES, midterm, final exams. Moreover, the LTA course was very theory-based (T10).

The same concerns and results could be observed in many research studies (Chung & Nam, 2018; Şahin, 2018; Vogt & Tsagari, 2014). As reported by these studies, many language teachers feel unprepared to assess their learners due to inadequate assessment training in pre-service teacher education programs. In some other studies (Hasselgreen, Carlsen & Helness, 2004; Volante & Fazio, 2007) it was revealed that teacher training programs do not provide adequate assessment literacy training for prospective teachers.

Considering certain needs in terms of knowledge, skills, and practice for LAL, the following comment can be considered as a common idea of several participants: “*I need more information on online testing (T2)*” and “*We can update ourselves with the changing world, like online assessment (T1)*”. As the views suggest, 42.8% of the interviewees (6 out of 14) self stated they need to learn about the latest trends such as online assessment.

Different from the others, some participants reported that they needed some training for improving validity in assessment practices noted as follows in their own words:

Especially, while designing tests validity can be a problematic area, therefore, I need some training in this area (T4).

I need to get further knowledge and practice in reliability and validity issues (...) in assessment (T5).

Moreover, with respect to designing a language test, some interviewees emphasized their need of knowledge related to creating test tasks and writing items which can be summarized as follows: “*I can say I need more training about writing test tasks and items and validation*” (T8).

When compared to similar studies in the related literature, it might be seen that there are certain alignments and discrepancies to some extent. For example, in a study conducted by Semiz and Odabaş (2016), it was found that EFL teachers need training on

“using portfolio, integrated language testing and assessment, placing students, etc.” while in this current study EFL teachers were generally found to need training on “Interpreting scores” and “Rating performance tests (speaking/writing)”. Moreover, the findings of this study were found in line with Hasselgreen, Carlsen, and Helness’s study (2004) which revealed that language teachers needed training in areas such as portfolio assessment, preparing classroom tests, peer and self-assessment, item writing, interviewing and rating among many other areas.

In sum, the in-service EFL teachers' opinions of their LAL needs are in line with the results emerging from the analyses of the interviews and the course documents. That’s to say, “Writing test tasks and items” and “validity” are the main topics generally needed by the interviewees even though they are covered in LTA courses.

When it comes to perceived LAL levels of in-service EFL teacher candidates, it is reasonable to interpret the data bearing the LAL needs in mind as it is anticipated that if the LAL level is good or excellent, then the LAL need should be in low or no need category. In order to discuss the levels of the in-service group of participants, Table 7 is prepared below.

Table 7

In-service EFL teachers' perceived LAL levels

LAL Items	Poor		Fair		Average		Good		Excellent	
	N	%	N	%	N	%	N	%	N	%
History of Language Testing	13	13.3	15	15.3	30	30.6	29	29.6	11	11.2
Procedures in language test design	2	2.0	7	7.1	27	27.6	39	39.8	23	23.5
Deciding what to test	1	1.0	1	1.0	9	9.2	49	50.0	38	38.8
Writing test specifications/blueprints	1	1.0	9	9.2	30	30.6	40	40.8	18	18.4
Writing test tasks and items			5	5.1	24	24.5	44	44.9	25	25.5
Evaluating language tests			3	3.1	17	17.3	49	50.0	29	29.6
Interpreting scores	1	1.0	3	3.1	13	13.3	51	52.0	30	30.6
Test analysis	1	1.0	8	8.2	25	25.5	43	43.9	21	21.4
Selecting tests for your own use			2	2.0	12	12.2	45	45.9	39	39.8
Reliability	1	1.0	1	1.0	21	21.4	52	53.1	23	23.5
Validation	1	1.0	1	1.0	18	18.4	53	54.1	25	25.5
Use of statistics	3	3.1	15	15.3	22	22.4	42	42.9	16	16.3
Rating performance tests (speaking/writing)			3	3.1	21	21.4	49	50.0	25	25.5
Scoring closed-response items			1	1.0	13	13.3	55	56.1	29	29.6
Classroom assessment					8	8.2	61	62.2	29	29.6
Large-scale testing			11	11.2	29	29.6	47	48.0	11	11.2
Standard setting			2	2.0	25	25.5	49	50.0	22	22.4
Preparing learners to take tests	1	1.0	2	2.0	11	11.2	59	60.2	25	25.5
Washback on the classroom	1	1.0	9	9.2	13	13.3	50	51.0	25	25.5
Test administration	1	1.0	3	3.1	15	15.3	43	43.9	36	36.7
Ethical considerations in testing	1	1.0	1	1.0	8	8.2	52	53.1	36	36.7
The uses of tests in society	3	3.1	10	10.2	33	33.7	42	42.9	10	10.2
Principles of educational measurement			7	7.1	30	30.6	45	45.9	16	16.3

The quantitative data above shows that in-service EFL teachers feel themselves competent in the following LAL topics with excellent level: “Selecting tests for your own use” (39.8%) and “Deciding what to test” (38.8%). Considering the items in detail, more than half of the participants ($n=53$) self stated that they have good level in terms of validity (54.1%) while 25.5 % reporting excellent level for the same item. The mostly mentioned items as in the poor and fair categories were found as “History of language testing” (28.6

%) and “use of statistics” (18.4%) which implies a need for further training. Furthermore, these items are followed with “large scale testing” and “the uses of tests in society” with the percentages of 11.2 and 13.3 respectively. Likewise, findings emerged from qualitative data obtained through interviews with in-service EFL teachers proves an alignment between the items which were labelled in a very high need for training category and low level of LAL. However, on the contrary to the findings aforementioned, Mede and Atay (2017) revealed a very limited level of LAL of Turkish EFL teachers particularly discussing deficiencies in test design and procedures, specific terms including validity, reliability in a mixed methods research study through questionnaire and focus group interviews. But, the in-service EFL teachers in this study self reported a good level for reliability and validity.

The overall perceptions about their LAL level, in-service EFL participants were given 5 options as “poor, fair, average, good, excellent” among which the highest mean values belonged to “deciding what to test” ($M=4.24$, $SD=.747$) and “ethical considerations in testing” ($M=4.23$, $SD=.729$) while the lowest mean value was reported as $M=3.10$ ($SD=1.197$) for the “history of language testing”. In a broad sense, the 51.0 % of in-service EFL teachers regarded themselves in “Good” level while almost equally as many respondents (48.0 %) considered their need of training as “Moderate”. This finding is in line with Kaya and Mede’s study (2021) which explored LAL levels of 195 EFL instructors reporting a satisfactory level of knowledge of assessment of participants while no evidence was found for the impact of certain factors such as experience and background on their LAL. However, the general tendency observed in the many research studies within the literature can be summarized as teachers’ having low assessment literacy levels (Hasselgreen, Carlsen & Helness, 2004; Tsagari & Vogt, 2017; Xu & Brown, 2016). For this reason, the qualitative data bears greatest importance as it provides deeper and richer insights for justifying the quantitative data.

In synthesis, Table 8 below shows the perceived LAL levels and needs of pre-service and in-service EFL teachers as a whole.

Table 8

A comparison of LAL needs and levels of all participants

Group of Participants	PERCEIVED LAL LEVELS					PERCEIVED LAL NEEDS				
	Poor	Fair	Average	Good	Excellent	No	Low	Moderate	High	Very high
Pre-service EFL teachers	2	4	36	45	9	2	21	47	21	5
	2.1%	4.1%	37.1%	46.4%	9.3%	2.1%	21.6%	48.5%	21.6%	5.2%
In-service EFL teachers	2	37	5	50	4	1	19	47	27	4
	2.0%	37.8%	5.1%	51.0%	4.1%	1.0%	19.4%	48.0%	27.6%	4.1%

As the results indicate, in-service EFL teachers regarded themselves in “good” category in terms of their LAL level with the highest percentage (51.0%). Similarly, pre-service EFL teachers self reported that they are in “good” level (46.4%), as well. However, on the contrary to this similarity, Mertler (2003) reported differences between pre-service and in-service teachers concluding that in-service teacher assessment literacy levels were significantly higher than the pre-service teachers’ level, which shows that classroom experience has a positive impact. Yet still, within this current study pre-service EFL teachers might feel prepared and competent enough as their LAL related knowledge has been just acquired in spring term when the data was collected. More specifically, as emphasized by DeLuca and Klinger (2010) assessment course had considerably positive impact on assessment literacy development of teacher candidates. On the other hand, from a wider perspective, as in-service EFL teachers’ level is slightly higher than pre-service teachers; one can conclude that teaching experience and real classroom practices are facilitating and mediating for LAL development.

Similar to these findings, the emergent qualitative data based on the analyses of the interviews with both in-service and pre-service EFL teachers and the course syllabus show that the majority of participants considered themselves in good or average level of LAL which contradicts with the aforementioned studies in the literature (Hasselgreen, Carlsen & Helness, 2004; Tsagari & Vogt, 2017; Xu & Brown, 2016) which mainly reported a low level of LAL.

In a more recent study, Yastibas and Takkac (2018) questioned the way language instructors improve LAL through interviews and discussions concluding that among three categories (previous assessment experience, assessment training and self-improvement) pre-service assessment training had the most influence on their development of LAL. In this regard, it is of great necessity to portray an overall picture of LAL levels and needs of pre-service and in-service EFL teachers in order to analyze the possible gaps between theory (knowledge acquired during pre-service teacher training) and practice (language assessment and testing practices in real classroom settings). Doing so, the language assessment literacy training in pre-service teacher education programs could be improved.

4.4. Results and Discussion for Research Question 3

What are the pre-service and in-service language teachers' beliefs about language assessment?

Concerning language assessment beliefs of the participants, some open ended questions were formulated and addressed in semi structured and focus groups interviews. When examined closely, the interpretations of qualitative data obtained from both pre-service EFL teachers and in-service EFL teachers yielded valuable results. As beliefs are difficult to define or show, the informants were asked different sub questions to unveil their beliefs and tendencies. To fulfil this aim, 12 questions were written as follows:

- 1) What are your reasons for assessing students?
- 2) How do you think assessment affect student learning?

- 3) How do you think assessment affect your teaching?
- 4) Can you name at least three useful assessment types-methods for student learning?
- 5) How do you make use of assessment results for students' improvement?
- 6) What do you think about the features of a good/ sound assessment?
- 7) What are the possible reasons for poor assessment?
- 8) What do you think of assessment for learning / assessment as learning / assessment of learning? Can you explain giving examples?
- 9) What do you think of students' responsibility (or role) in assessment process?
- 10) What does "language assessment literacy" mean for you?
- 11) Of the topics above (in the questionnaire), which one / ones do you think the most helpful to language teachers?
- 12) Of the topics above (in the questionnaire), which one / ones do you think the most challenging for language teachers?

Therefore, each interview question will be dealt with one by one for each group of participants under related research questions. In this sense, the first ten questions will be discussed under the third research question with regards to language assessment beliefs while the last two questions will be analyzed under the fifth and the sixth research questions which discuss the importance level of language assessment related topics. While reporting the results, certain interview questions were grouped as they point out similar themes and issues.

4.4.1. Results and Discussion for Assessment Beliefs

4.4.2. Results and Discussion for Interview Question 1 and Question 9

What are your reasons for assessing students? (IQ1) What do you think of students' responsibility (or role) in assessment process? (IQ9)

In this part, two interview questions will be discussed in order to identify participants' beliefs about the reasons for assessment and the role of students in assessment. The qualitative data emerged from the interviews were analysed using content analysis method through deductive and inductive approach. Concerning the common themes and similar ideas with regards to the main reasons for assessing students, the following categories in Table 9 and Table 10 were constructed.

The aim of Interview Question 1 was to reveal the perceptions of EFL teachers and teacher candidates with regard to the purposes of assessment, that's the reasons for assessing students. The key results with respect to the opinions shared by pre-service EFL teacher candidates are displayed in Table 9 below.

Table 9

Main findings identified based on pre-service EFL teachers' responses in IQ1

Interviewee Codes	Category of Reasons	Description of Main Themes	Examples drawn from the interviews
S1, S2	Assessment of learning	Make decision	S1: To understand how much they have learnt from instruction S2: Generally, we use assessment whether they fail or pass.
S2	Assessment for learning	Improvement	S2: (...) to educate and improve the skills of individuals so as to help them be ready for their future and occupation.
S1		Diagnosis	S1: To understand which parts they have learnt or not, thus arranging future lessons
S1	Assessment as learning	Reflecting	S1: To see what necessary precautions to take during the teaching

As the findings indicate, the main themes identified in this question with regards to the main purposes of language assessment were grouped as follows: assessment for

learning, assessment as learning and assessment of learning. In each group of category, some themes were determined based on a cluster of responses. Therefore, data yielded from IQ1 showed that making decision and diagnosing purposes were mostly favoured ones by interviewees seen in the following quotation as well:

To understand how much they have learnt from instruction (S1).

Concerned with the category of assessment for learning, the general tendency of interviewees was found to address “improvement” and “diagnosis” in relation to the purpose of assessment which revealed their wish to foster learning and improve the skills of learners. This perspective can be clearly seen in the following utterance of interviewee S2:

The main purpose of language assessment is to educate and improve the skills of individuals so as to help them be ready for their future and occupation.

As the Table 9 displayed, “assessment as learning” can be characterised by interviewee wish to reflect on their teaching and student learning, so that they can arrange and design their future lessons. Considering the mostly mentioned themes which arise from interviewee comments, summarised above and exemplified as well, “diagnosing” and “making decision” suggest that the pre-service teacher interviewees in this study attempt to practice language assessment for instructional purposes, so that they can design their lessons.

The main purposes of language assessment pointed out by in-service EFL teachers were summarized and exemplified below in Table 10.

Table 10

Main findings identified based on in-service EFL teachers responses in IQ1

Interviewee Codes	Category of Reasons	Description of Main Themes	Examples drawn from the interviews
T3,T4, T5, T6, T8, T10, T14	Assessment of learning	Make decision	T3: To understand how much they have learnt from instruction T8: ... assessment aims at providing evidence for judgements of the assessors about the performance, proficiency, and/or aptitude of a learner
T3 ,T7, T10		Accountability	T3: To evaluate the success of the instruction T7: I assess students in order to measure the effectiveness of the course
T1,T2, T5,T8,T10	Assessment for learning	Improvement	T2: to improve the quality and effectiveness of education T5: to see their developmental process
T1,T2,T3,T4,T5,T9, T11, T13		Diagnosis	T2: To be able to identify the problems regarding the teaching and learning process T9: to diagnose their lacks in learning process
T1,T3,T5,T6,T8,T9, T11, T12	Assessment as learning	Reflecting	T1: to self-criticize my own teaching T6: to make them become aware of their own performance
T3,T4, T5, T10		Motivation	T4: To encourage them to study more T3: To motivate students to study

The main themes identified in Table 10 above, drawn from responses obtained through RQ3, associate to interviewees' opinions on the main reasons for language assessment which were categorized under the following headings, that's assessment as learning, assessment of learning and assessment for learning. Considering the mostly mentioned themes which arise from interviewee comments, summarised above and exemplified as well, "make decision" and "accountability" were evaluated as "assessment of learning" because of the fact that evaluating the end product of education implies product based assessment. For example, Teacher 5 commented in the following way "to what extent they have learnt the points that we covered in the lesson" as one of the reasons for assessment. Similarly, the successive idea was shared by Teacher 3 as "To understand how much they have learnt from instruction".

With respect to assessment for learning, the general tendency of interviewees was found to address "improvement" and "diagnosis" in relation to the purpose of assessment.

To move discussion further, these themes suggest that in-service EFL teachers consider that the main purposes of assessment should be for enhancing the quality of education and teaching instruction, identifying the problems, strengths and weaknesses within the teaching and learning process. The following comments by Teacher 8 and Teacher 5 reflect similar views on the mentioned reasons as follows:

“Assessment aims to guide, facilitate, and accelerate learning process” (T8)

“To diagnose their strengths and weaknesses” (T5)

When analyzed closely, “assessment as learning” can be characterised by interviewee wish to reflect on their teaching and student learning with an attempt to increase motivation in the long term. When evaluating related themes summarised in Table 10 above having high frequency under “assessment as learning” perspective, it can be noted that motivation and reflecting purposes were self reported by the in-service EFL teachers during interviews, which implies that learner based assessment was embraced by the majority of the EFL teachers within the study. To illustrate, some views can be emphasized in their own words as in the following utterances:

“The main reasons of assessment are to self-criticize my own teaching and organize the lessons accordingly” (T1)

“To make them become aware of their own performance” (T6)

The main themes constructed based on the responses of participants with respect to the purposes of language assessment were found in alignment with Fulcher’s comments (2010). That’s, “improvement, motivation, feedback” as mentioned by participants were also supported by the statement of Fulcher which is noted as “Tests encourage learning because they are gateways to goals (...) not only by motivating learners, but also providing feedback on learning and achievement to both learners and teachers” (2010: 67). Therefore, aligning teaching and learning objectives, instructions, and outcomes well with assessment is crucial as it ensures a meaningful reason for assessment, thus justifying the purpose. Moreover, “internal” tests generally prepared by

teachers themselves for classroom use, thus considered learning-oriented assessment, have been used for diagnostic purposes, such as exploring the extent to which they have learnt, the difficulties they might face, the strengths and weaknesses each learner has (Fulcher, 2010; Hamp-Lyons, 2016). This general tendency is also embraced by the participant teachers within this present study characterized by the following comments: “To understand how much they have learnt from instruction” (T3), “to diagnose their strengths and weaknesses” (T9), “to what extent they've learned the points that we covered in the lessons” (T5) and “to find out whether the objectives are achieved or not” (T14).

Considering the conceptualisation of the role of students in the language assessment process, each group of participants were asked the same question during the interviews. The results are displayed below reporting the findings belonged to both pre-service and in-service EFL teachers.

Table 11

Perspectives of participants for student responsibility in language assessment

Group of Participants	Category	Main Themes	N	%	Example quotations
Pre-service EFL teachers	Active Role	○ Have an active role	2	100%	S1: They should present what they know S2: An active role showing what they know
	Reflective Role	○ Feedback provider	2	100%	S1: provides feedback for the teacher S2: Reflecting the feelings such as being disappointed or not with the results
	Proactive Role	○ Questioning	1	50%	S2: (...) questions the content and types of tasks in the exam, check expectations
In-service EFL teachers	Active Role	○ Take responsibility ○ Have an active role ○ Being autonomous	13	92.8%	T1: taking responsibility T4: They need to take active role in assessment
	Reflective Role	○ Judge themselves ○ Awareness of their own learning ○ Self assessment	7	50%	T2: They should learn not only their grades but objectively judge how and why they got that result. T5: Keep track of his/her performance through self-evaluation T8: When students are involved in the assessment process, assessment not only becomes more inclusive but also enhances students' awareness of their own learning and helps them become more autonomous learners
	Collaborative-Interactive Role	○ Peer assessment ○ Ask for support	5	35.7%	T2: if they have any ideas on how to be more successful, they can collaborate with their teachers T4: they can give feedback to their peers. T5: Ask for support when necessary
	Proactive Role	○ Get ready for assessment ○ Familiarizing with the instruction, criteria, objectives	2	14.2%	T2: I mean they should do their share of the work firstly by getting ready for the assessment, paying attention to criteria and objectives, familiarizing themselves with the instructions, etc.

Consideration of the results, the interview question 9 portrays interviewees' opinions on the students' responsibility in language assessment process. Thus, information drawn from this interview question, summarised in Table 11 above, is concerned with the role of students in language assessment process categorized as active, proactive, reflective and collaborative-interactive by the researcher. As the results demonstrate, it can be observed that interviewees generally are of the opinion that student should be given an active role in assessment.

When in-service EFL teacher data is analyzed in a detailed manner, it can be seen that the highest percentage belongs to the following categories with respect to student responsibility: “active” (92.8%) and “reflective” (50%). Other key categories, on the other hand include “collaborative-interactive” and “proactive” with the percentages 35.7% and 14.2% respectively. What’s more, the evidence collected from pre-service data revealed the following roles of learners in language assessment process: active role, reflective role by feedback providing and proactive role by questioning.

In assessment process, student responsibility bears greatest importance as they are responsible for monitoring their own learning which is also supported by the interviewees in this study. The themes formulated based on the responses of informants with respect to the role of students were categorized as active, proactive, reflective and collaborative-interactive each of which has various themes. Initially, as an active role, the students are expected to self-regulate their own assessment process, thus becoming autonomous learners. In similar vein, as noted by Spiller (2012), the responsibility of students requires to manage “assessment design, choices, criteria and making judgments”. In proactive role, the students generally may pay attention to the criteria and goals, and familiarize themselves with the instruction which is also voiced by Teacher 2 and Student 2 as follows in their own words:

I mean they should do their share of the work firstly by getting ready for the assessment, paying attention to criteria and objectives, familiarizing themselves with the instructions, etc. (T2).

I think this can be explained by introduction, body and conclusion roles. Before the exam, the student questions the content and types of tasks in the exam, check expectations (S2).

Moreover, reflective role necessitates monitoring the progress of their own learning

and assessment through self assessment in addition to making decisions and judgments of their performance which is shown in the following opinions of Teacher 2 and Teacher 5:

They should learn not only their grades but objectively judge how and why they got that result (T2).

Keep track of his/her performance through self assessment (T5).

Finally, collaborative-interactive role of students in assessment process implies that they give feedback to their peers and ask for help when necessary. In synthesis, at different stages of assessment practices; active, reflective and collaborative-interactive roles enhance students learning in which interaction and analyzing their own or peers' works provide and necessitate a critical point of view.

4.4.3. Results and Discussion for Interview Question 2, Question 3 and Question 5

How do you think assessment affect student learning? (IQ2), How do you think assessment affect your teaching? (IQ3), and How do you make use of assessment results for students' improvement? (IQ5)

In this part, three interview questions (2, 3, and 5) will be discussed together as they imply the effects of language assessment on students or teachers. In this regard, it was intended that Interview Question 2 would provide information connected to the effect of language assessment on students learning based on the comments of interviewees. The following Table 12 attempts to summarise and exemplify the main findings including mostly mentioned themes, example comments and categories as well. Therefore, the qualitative data collected from this interview question was arranged below.

Table 12

Main themes identified in IQ2 based on pre-service EFL teachers responses

Interviewee Codes	Category of Impact	Description of Main Themes	Examples drawn from the interviews
S1, S2	Positive	Motivation	S1: It motivates students
S1	Positive	Feedback	S1: It provides feedback for themselves about what they have learnt or not
S2	Positive	Diagnosis	S2: To what extent they have learnt, the missing partsd they can see
S1, S2	Negative	Stress	S2: As they want to pass the exam, they generally feel stressed S1: Because of the exam oriented system, the feel stressed.

Associated with the effect of language assessment on student learning, the main information drawn from the table above indicates that the first theme in positive category and only one negative theme stood out as having the highest percentages based on the interviewees' responses. These themes were regarded as "motivation" and "stress" which reveal the importance attached to washback effect by interviewees either in positive or negative way.

When in-service EFL teachers' opinions are concerned, the following table intends to portray the issue discussing identified themes either positive or negative.

Table 13

Main themes identified in IQ2 based on in-service EFL teachers responses

Interviewee Codes	Category of Impact	Description of Main Themes	Examples drawn from the interviews
T1, T3,T4,T5, T11	Positive	Motivation	T3: It motivates students for studying T4: I think it motivates learners and create a reason to study more
T2, T5, T8, T10, T11	Positive	Awareness	T8: Assessment raises the awareness of the learners both about their own performance.....
T5,T8,T9,T11, T14	Positive	Feedback	T5: it provides feedback for their weaknesses and strengths
T1, T2, T9, T10, T13	Positive	Diagnosis	T2: help students realize the weaknesses
T1, T2, T5, T10	Positive	Improvement	T1: improve their own skills T2: they can compensate for the missing information, practice more and take one more step towards being better
T1, T2, T4, T11, T12	Negative	Stress	T11: As the system is exam-oriented, they generally feel stressed before exams T4: sometimes, it becomes competitive and this feeling of competing may affect learners negative way

With relate to the effect of language assessment on students learning, the Table 13 above demonstrates the mostly mentioned themes categorized as positive or negative. Consideration of the results, it can be stated that the majority of the in-service EFL interviewees embody positive sides of assessment on learners with a very high level of percentage (83.3%). In this sense, regarding the positive effect of language assessment, “motivation”, “awareness” and “feedback” were found as recurrent themes notified by the interviewees with 35%. As a general picture, the results indicate that the interviewees mostly have a tendency towards positive impact of language assessment on students by expressing the following effects such as creating awareness, providing feedback, realizing weaknesses and strengths, increasing motivation and participation, improving learning, among others. Studies such as those conducted by Choi (2008) and Li, Zhong, and Suen (2012) agree with the findings of this study. To illustrate, Li et al. (2012) carried out a study about a large scale test (CET in China) concluding that the test had an impact on arranging the content of the course in addition to the time devoted for each topic.

Moreover, Choi (2008) supported the view that language testing plays a big role in the objectives, approach and motivation for language learning.

For instance, with respect to “awareness”, T8 acknowledged that “assessment raises the awareness of the learners both about their own performance and the expectations from them”. Moreover the following comment can be noted as an illustrative opinion of a considerable number of participants with related to motivation: “Assessment helps learners practice more and take one more step towards being better” (T2).

On the other hand, as it is displayed above, the main negative impact of language assessment on students’ learning was considered as “stress” with a 28.5% percentage which can be exemplified in the following opinion of T4: “But sometimes it becomes competitive and this feeling of competing may affect learners in a negative way”. Thus, being tested might create an emotional barrier for them during their language learning process by resulting in stress, anxiety, fear of failure, a decrease in success, among others. To put it short, as reported by the majority of interviewees (both pre-service and in-service EFL teachers), stress is the main negative unintended impact of language tests which rises learners’ anxiety levels thus hindering their success (Fulcher, 2010).

The main purpose of the interview question 3 was to ascertain the effect of language assessment on teachers’ teaching practices based on the responses provided by interviewees. Thus, the Table 14 was designed in order to display the main findings and common themes.

Table 14

Main themes identified in IQ3 based on in-service EFL teachers responses

Interviewee Codes	Category	Description of Main Themes	Examples drawn from the interviews
T1, T2, T3, T7, T8, T11, T14	As Teaching	Diagnosing (Identifying Lacks)	T3: S/he can see both her/his and the students' strengths and weaknesses
T1, T2, T3, T7, T8, T11, T12	After Teaching	Organizing and Arranging Lessons	T1: I can prepare my materials accordingly
T2, T8, T11, T12	After Teaching	Critical Thinking and Reflecting	T2: I use assessment to take lessons about what I teach and how I teach T8: give me feedback on my own teaching performance.

Concerning with the effect of language assessment on teaching practices, two main categories were constructed based on the similar opinions of participants with respect to the question, to name “As teaching” and “After teaching”. Therefore, the main information drawn from the table above indicates that theme 1 stood out as having the highest percentages based on the interviewees’ responses under the first category named “as teaching”. This theme was regarded as “Diagnosing (Identifying Lacks)” (50%) which reveals the importance attached to explorative aspect of language assessment impacts by interviewees. “Organizing and Arranging” and “Critical thinking and Reflecting” were found the mostly mentioned themes under the second category named as “After teaching” with relatively large percentages, 50% and 28.5% respectively. For instance, the opinion of Teacher 7 highlights the impacts of language assessment with regards to both diagnosing and organizing concerns: “I most of the time do remedy teaching if I see that there are topics that are not covered by most of the students”. As it can be seen here, the teacher makes use of the assessment by identifying the missing topics and arranging her lessons accordingly. The conclusions of this study highlighted above are in line with Winke’s study (2011) which reveals that the test (English Language Proficiency Assessment in his case) enabled the teachers to identify learning needs of students, thus designing learning materials suitable for their needs.

With regards to the pre-service EFL teachers' opinions for the same question, the following table is prepared.

Table 15

Main themes identified in IQ3 based on pre-service EFL teachers responses

Interviewee Codes	Category	Description of Main Themes	Examples drawn from the interviews
S1, S2	As Teaching	Diagnosing (Identifying Lacks)	S1: I see their current situation, identifying problems, missing parts, etc.
S1	After Teaching	Organizing and Arranging Lessons	S1: I can design my future lessons by making necessary arrangements
S2	After Teaching	Reflecting	S2: What are the missing parts, so I can see the effectiveness of my teaching practice
S2	After Teaching	A decrease in job satisfaction	S2: (...) I have to design such tests I do not like personally which in turn decreases my job satisfaction.

When it comes to the impact of language assessment on teaching practice, two main categories were named based on the recurrent ideas of participants as follows “Impacts as teaching” and “Impacts after teaching”. To illustrate, the main information drawn from the Table 15 above indicates that diagnosing effect may be observed as teaching which enables identifying lacks and missing parts in teaching and learning process. In “after teaching” category, three main themes stood out as model responses which were coded as “organizing and arranging lessons”, “critical thinking and reflecting” and “ a decrease in job satisfaction”. For instance, the following comment clarifies the reason for the decrease in job satisfaction:

Because of the expectations of the system, school administration, and parents, I have to design such tests I do not like personally which in turn decreases my job satisfaction (S2).

In sum, with regards to the impacts of a language test, the opinions of interviewees in general sense agree with Fulcher's perspectives in which he comprehensively pinpointed the following effects a test has as follows:

- what teachers teach
- how teachers teach
- what learners learn
- how learners learn
- the rate and sequence of teaching
- the rate and sequence of learning
- attitudes to the content, method, etc. of teaching and learning (Fulcher, 2010: 277).

It was hoped that interview question 5 associated with using assessment results would help to identify the opinions of in-service and pre-service EFL teachers about the issue. To fulfil this aim, the following table is drawn summarizing the results and the main themes.

Table 16

Main themes identified in IQ5 based on in-service EFL teachers responses

Interviewee Codes	Use of Assessment Results	Description of Main Themes	Examples drawn from the interviews
T2, T3, T4, T5, T8, T9, T11, T12, T13, T14	Learning	Providing feedback	T4: By giving feedback on their assessment results
T1, T2, T3, T5, T7, T8	Instructional	Organizing lessons	T5: I design lesson activities accordingly
T1, T3, T10	Instructional	Exploring the lacks	T1: I detect the topics they are lack of learning
T2, T8, T11, T14	Reflecting	Discussing the reasons for failure	T2: I ask them about the reasons of their failures. Based on their feedback, we discuss what should be done, try to reach mutual understanding T8: Also, I use the results to monitor and improve my own practice of the program

When it comes to the use of assessment results, three main categories were named based on the recurrent ideas of participants as follows “Learning”, “Instructional” and “Reflecting”. To illustrate, among the uses of assessment results, 71.4 % of the interviewees identified “providing feedback” as the main theme under “Learning” category, which reveals the importance attached to making use of assessment results for constructive feedback by interviewees. This point of view can be easily seen in the following comment:

I use assessment results to give feedback to my students on their performance and their weak and strong points. And consequently, I use the results to guide the students about how they can improve their learning (T8).

Moreover, the main information drawn from the table above indicates that using assessment results for instructional purposes stood out as having the highest percentages based on the interviewees’ responses including the following themes “Organizing lessons” and “Exploring the lacks” with 42.8% and 21.4% percentages respectively.

Another recurrent theme was observed in one of the interviewee response which is named as “reflecting”, as a matter of fact that the learners are expected to think over their results and criticize their own failure on the basis of the interviewee comment below:

I ask them about the reasons of their failures. Based on their feedback, we discuss what should be done, try to reach mutual understanding and plan our upcoming lessons accordingly (T2)

With regards to the pre-service EFL teachers’ opinions for the same question, the following table is prepared.

Table 17

Main themes identified in IQ5 based on pre-service EFL teachers responses

Interviewee Codes	Use of Assessment Results	Description of Main Themes	Examples drawn from the interviews
S2	Learning	Providing feedback	S2: As the results suggest summative assessment, I can use them in a formative way by providing feedback.
S1, S2	Instructional	Organizing lessons	S1: (...) I decide on the difficulty level of the next test and arrange it accordingly
S1, S2	Instructional	Exploring the lacks	S2: I detect the topics they haven’t learnt
S2	Reflecting	Tailor the teaching	S2: I can decide about the things I need to change, tailor about my teaching

When it comes to the use of assessment results, the obtained pre-service EFL teachers’ data revealed that they plan to use assessment results for learning, instructional and reflective purposes. In doing so, learning purpose indicates that the main purpose of getting assessment results is to provide feedback according to the interviewees. Moreover, for the second category, the main information displayed in Table 17 uncovered two main themes, to name organizing lessons and exploring lacks which can be exemplified in the following comment by Student 1:

After using an assessment task, I look at the results and understand the difficulty of test, so I decide on the difficulty level of the next test and arrange it accordingly.

When evaluating other themes summarised in Table 17 above having high frequency, it can be noted that “tailoring teaching” is another recurrent theme which requires reflective thinking on the way of one’s own teaching.

4.4.4. Results and Discussion for Interview Question 6 and Question 7

What do you think about the features of a good/ sound assessment? (IQ6) What are the possible reasons for poor assessment? (IQ7)

These two interview questions were posed in order to determine the basic qualities of a good or poor assessment practice based on the respondents’ perceptions. Therefore, interview question 6 aims at eliciting the answers about the basic tenets and features of a good assessment which were summarized below in Table 18.

Table 18

Main themes identified in IQ6 based on in-service EFL teachers responses

Interviewee Codes	Percent	Description of Main Themes	Examples drawn from the interviews
T1, T2, T4, T5, T7, T8, T10, T11, T14	64.2%	Valid	T2: It should assess what it was planned to test (the assessment tool should cover the objectives of the course), and the assessment results should be consistent
T1, T2, T5, T8, T10, T11	42.8%	Reliable	T10: A good assessment should be valid and reliable in the first place.
T4, T5, T9, T13	28.5%	Process oriented	T4: it should be process based. I think assessment should provide opportunities to reflect their development over time
T5, T8, T11	21.4%	Positive washback	T10: it should create a positive wash-back effect
T5, T6, T8, T11	28.5%	Providing feedback	T6: A good assessment should inform teaching and learning
T10, T14	14.2%	Authentic	T14: A good language assessment should be related to real life.

Concerning the key considerations for a sound and good assessment, in-service EFL teachers pinpointed the following themes as “valid, reliable, process-oriented, authentic, having positive washback effect, providing feedback”. When analyzed in detail, all interviewees ($n=14$) supported that a sound language assessment need to be valid in terms of its content, purpose, format, etc. Similarly, these basic features such as reliability and validity are among the top qualities for a language assessment task to have a desired impact on learning and teaching. In similar vein, the qualities of a good language assessment mentioned by the participants within this current study agree with those reported by Giraldo (2019: 56) as “valid, reliable, sensitive to students’ affect and that it provides feedback to improve learning”. Moreover, two interviewees suggested that a language assessment needs to be objective (T9) and friendly (T1).

In addition to these themes, a core list of basic features of a good language test can be more comprehensively summarized as follows: valid, reliable, practical, comprehensive, relevant, balanced, clear, authentic, practical, objective (without rater’s bias), appropriate, and etc. which can be categorized as test-related, administration-related and teacher-related qualities (elttguide.com, 2020; Paradowski, 2002).

With regards to the pre-service EFL teachers’ opinions for the same question, the following table is prepared.

Table 19

Main themes identified in IQ6 based on pre-service EFL teachers responses

Interviewee Codes	Description of Main Themes	Examples drawn from the interviews
S1, S2	Valid	S2: The student should know what topics are included in the test
S1, S2	Reliable	S1: Needless to say, a good assessment should be valid and reliable.
S1, S2	Providing feedback	S1: The student should benefit from the results instead of learning the score.
S2	Authentic	S2: It should be authentic so the results would be meaningful.

Considering the qualities of a good language test, pre-service EFL teacher data seemed in alignment with in-service teacher data. Thus, a good way of language assessment should address certain considerations which are summarized as reliability, validity, authenticity and feedback providing by the interviewees. The findings of Hatipoğlu's study (2010) agree with these findings in that "*the testing language skills/knowledge, reliability and validity*" were the most reported topics by the participants who claimed that these mentioned topics would ensure them to "*prepare fair tests*" within Hatipoğlu's study (2010).

More importantly, one pre-service teacher candidate emphasized the benefit the test taker should have from a test uttered as in the following way:

The student should benefit from the results instead of learning the score (S1).

The aim of Interview Question 7 was to unveil the perceptions of EFL teachers and teacher candidates with respect to the basic features of poor assessment practices. The main results related to the opinions shared by in-service EFL teachers are displayed in Table 20 below.

Table 20

Main themes identified in IQ7 based on in-service EFL teacher responses

Interviewee Codes	Category of Qualities	Description of Main Themes	Examples drawn from the interviews
T1, T2, T3, T5, T8, T10, T11	Test-related	Lack of validity	T3: Items not covering the instruction
T1, T2, T5, T11	Test-related	Lack of reliability	T1: If the test isn't valid and reliable
T3, T5,	Test-related	Poor items	T3: Poor items (Not understandable, vague, not assessing what it is meant to assess)
T2, T5, T11	Teacher-related	Lack of a purpose	T2: I also believe in the necessity of teaching students about the purpose of assessment and how it is going to be conducted
T4,T5,	Teacher-related	Lack of knowledge of teachers about assessment type, method, criteria, etc.	T4: Teachers' lack of knowledge about different assessment methods

Regarding the features of a poor language assessment, two main categories were constructed as test-related and teacher-related problems. Concerned with the test related problems, the main reasons were mentioned as “lack of validity (50%), lack of reliability (28.5), lack of purpose (21.4%) and poor items (14.2%)”. The opposite of these reasons were also mentioned as the qualities of a good language assessment earlier. Also, the majority of interviewees mentioned test related concerns exemplified as follows:

Not determining test specifications or poorly determined ones (T5).

Items not covering the instruction or not letting the teacher actually evaluate the students (T3).

Teacher-oriented problems were generally considered as the lack of knowledge with regards to assessment methods, techniques, criteria, etc. The following comment by Teacher 8 puts a great emphasis on the importance of the teacher (as assessor or rater) for the quality of an assessment procedure from the designing phase to interpreting the results phase:

Training of the raters and proctors may be missing or of poor-standard, so the results may not reflect the learners' real performance even if the assessment tools are relatively well-designed (T8).

As qualitatively searched, poor assessment practices and their consequences are detailed in a study reporting “practicality, learners, test design, administrative obligations and teachers” are among the main causes of challenges (Ashraf & Zaki, 2019). Moreover, in addition to inadequate assessment literacy of teachers; class size is another problem resulting in unfruitful use of materials, time-consuming nature of alternative assessment methods, unmanageable workload, among others. Moreover, without considering the context and teaching goals, pre designed appointed assessment tasks would produce inadequate and even deceptive information about students learning thus resulting in unfair decisions (Koh, et al. 2018). In similar vein, as supported by Fulcher (2010), it is quite necessary to take social, ethical and historical context into consideration to produce good language tests. In this sense, classroom as a context bears greatest importance which might yield a wide range of functions and meanings of assessment for students and teachers as a test taker, test developer and / or as an assessor. When the context sensitive knowledge is not taken into account, certain problems might occur about the content of language assessment (test items problems, difficulty level, authenticity, face and content validity concerns, etc.) and the impact of language assessment (positive or negative washback, disappointments, etc.) can be a serious concern.

With regards to the pre-service EFL teachers' opinions for the same question, the following table is prepared.

Table 21

Main themes identified in IQ7 based on pre-service EFL teachers responses

Interviewee Codes	Category of Qualities	Description of Main Themes	Examples drawn from the interviews
S1, S2	Test-related	Lack of validity	S1: not assessing what it is meant to assess
S1, S2	Test-related	Lack of reliability	S2: If the test isn't valid and reliable
S1	Test-related	Poor items	S1: If the test is poorly designed
S2	Teacher-related	Not authentic	S2: Using the same questions and test types over and over again which reduces reliability and validity as well

Associated with the qualities of a poor language assessment, the pre-service EFL teacher data supports the previous data drawn from in-service EFL teachers. In this regard, poor language assessment may derive from lack of reliability, validity, authenticity along with poor test items. These qualities may be caused by either the test itself or the teacher herself / himself as categorized and discussed before.

4.4.5. Results and Discussion for Interview Question 4 and Question 8

Can you name at least three useful assessment types-methods for student learning? (IQ4) What do you think of assessment for / as / of learning? Can you explain giving examples? (IQ8)

Interview questions 4 and 8 are grouped as they deal with the assessment approaches and methods. By doing so, the main approach adopted by interviewees and the most useful assessment types will be discussed below.

In order to identify the most useful language assessment methods-types, the interviewees were invited to answer the fourth question the results of which are reported and summarized below in Table 22.

Table 22

Main themes identified in IQ4 based on all participants' responses

Group of Participants	The most useful assessment types	N	%	Interviewee Codes
Pre-service EFL teachers	Formative Assessment	2	100%	S1,S2
	Portfolio Assessment	2	100%	S1,S2
	Skill-based Assessment	1	50%	S1
	Alternative Assessment	2	100%	S1,S2
In-service EFL teachers	Formative Assessment	7	50%	T1,T2,T4,T6,T9,T10,T11,T13
	Portfolio Assessment	6	42.8%	T1,T3,T8,T9,T10, T14
	Written-Oral exams (assignments, discussions)	4	28.5%	T1,T3, T4, T7
	Alternative Assessment	3	21.4%	T5, T9, T11
	Diagnostic assessment	3	21.4%	T2, T5, T6

In pursuit of answering the fourth interview question related to useful assessment methods, the qualitative data concurrently reported with quotes were analyzed and interpreted. To start, formative assessment was found as the most useful assessment method based on the responses of the majority of informants (56%) as supported and exemplified in the following comment of one of the interviewees:

Language learning is a process which requires discipline, regular practice and sustained motivation and formative assessment is a great tool for both being aware of student progress and act accordingly and give students a reason for developing self-regulation strategies and to motivate them throughout this process (T2).

Meanwhile, one of the interviewee (T8) also noted that self assessment and dynamic assessment are the ones which she finds very useful. To support these views, Teacher 2 explains the reason why she advocated for diagnostic assessment:

“The first assessment type that I make use of is the diagnostic one because without knowing the current level of the students, it does not seem likely to achieve success” (T2).

However, the in-service EFL teachers’ perceptions about useful assessment tools show that they self reported the positive sides of formative assessment, alternative assessment, portfolio assessment and diagnostic assessment; yet, still most of them do not allot much time to using them. This might result from their previous experiences related to assessment, inadequate time, class size, limited knowledge for alternative assessment tasks, heavy workload, etc. Therefore, these findings of this study agree with Büyükkarcı’s study which reported positive attitudes of in-service EFL teachers to formative assessment (2014). However, the real classroom practices do not reflect these positive perspectives as a result of large class sizes and workload. In this sense, as previously noted, both contextual and experiential factors shape assessment practices of language teachers by influencing their tendency to use a certain type or tool of an assessment as they mostly prefer to use familiar assessment methods they have experienced as test takers (Reynolds-Keefers, 2010).

These results also have a lot in common with the study of Rogers et al. (2007) which also concluded that non traditional assessment (alternative assessment and formative assessment) were advocated by participants, but still it was observed application of pen and paper tests. In similar vein, Davidson and Coombe (2019) explained the negative assessment related experiences of teachers as test takers in a detailed manner mentioning difficult tests, unawareness of the content, question formats, or time of a test, poorly written questions, insufficient time, rater’s bias, fruitless communication of test scores, etc. which in long term shapes the way they assess their own learners.

With an attempt to explore the perceptions of participants about “assessment as learning, assessment for learning and assessment of learning” issues, the interviewees were posed the eighth interview question the results of which are explained and discussed below. As for the eighth open ended question in the interview, the majority of the in-

service EFL teachers emphasized “assessment for learning” as they thought this type of assessment puts the student and her / his needs at the centre. They stated that by doing so they enhance student learning and teaching them how to assess themselves and be responsible for their own learning. To support their view, two of the interviewees acknowledged that:

“Assessment for learning focuses on students and their learning, it aims to improve learner performance” (T6).

“Assessment for learning refers to use of assessment results by teachers to give feedback to students, inform their future instructional decisions, and improve their teaching practices to increase the learners’ performance” (T8).

A point worth noticing is that formative assessment entails a collaborative interaction between the learner and the teacher, thus creating an engaging and supportive atmosphere during the language assessment process. In this sense, Giraldo (2021) showed that an appropriate assessment based on LAL lead to a positive change and make a great deal of contribution to students learning and the program itself. Remarkably, assessment for learning or learning-oriented assessment in the form of formative assessment, self and peer assessment, etc. attempts to enhance student learning on the contrary to scoring or certification. Similarly, through assessment activities the learner might diagnose her/ his weaknesses and strengths, thus improve one’s own learning (Hamp-Lyons, 2016).

4.5. Results and Discussion for Research Question 4

What is the contribution of language testing and assessment (LTA) course offered in pre-service language teacher education programs? (RQ4)

To answer this research question one open ended question is placed in the questionnaire as “*Have you ever taken a training /course/ professional development program on language testing and assessment?*” with three options as “No Training”, “A

Little Training”, and “Extensive Training”. Further one open ended question is posed in the interview session as “*Do you think language assessment course in pre service teacher education has met your needs of assessment? If yes, how? If not, why not?*” with an attempt to question the general situation of the Language and Assessment course based on the informants’ responses.

To fulfil these purposes, the following table is prepared to display and summarize the results.

Table 23

A summary of training, LAL level, needs and language assessment course views of participants

Research Foci	Category	Pre-Service Teachers		In-Service Teachers	
		N	%	N	%
Training	No	4	4.1	10	10.2
	A Little	44	45.4	67	68.4
	Extensive	48	49.5	21	21.4
Perceived LAL Level	Poor	2	2.1	2	2.0
	Fair	4	4.1	37	37.8
	Average	36	37.1	5	5.1
	Good	45	46.4	50	51.0
	Excellent	9	9.3	4	4.1
Perceived LAL Need	No	2	2.1	1	1.0
	Low	21	21.6	19	19.4
	Moderate	47	48.5	47	48.0
	High	21	21.6	27	27.6
	Very High	5	5.2	4	4.1
LTA course to meet LAL needs	Yes	1	50	3	21.4
	No	-	-	7	50
	To some extent	1	50	4	28.5

As the table above portrays; the results for the training they have received, their perceived LAL levels and further training needs are summarized. According to the results, the question related to their received trainings uncovered that 49.5 % of pre-service EFL teachers self stated their having extensive training while the majority of in-service EFL teachers (67 out of 98) reported that they have had a little training (68.4 %). When their LAL levels are concerned, the same picture is witnessed for each group of participants, that's to say "Good" level was the mostly mentioned LAL level of the 46.4 % of pre-service EFL teachers and 51.0% of in-service EFL teachers. Concerning further training for LAL development, the general tendency of each group of participants again reflects the same results. To illustrate, 68 of the pre-service EFL teachers (70.1%) and 74 of in-service EFL teachers (75.6%) considered their needs either moderate or high.

However, 4 out of 14 in-service EFL teachers are of the opinion that even though they have training to some extent at their pre-service teacher training, they generally criticize the LTA course being too theoretical, that's leaving no room for practice. What's more, half of the interviewees (7 out of 14) argue that the LTA course has had no facilitative or beneficial effect on their LAL development which raises serious concerns for the design and implementation of the course. These differing points of views either positive or negative can be observed in the following quotations:

In terms of theoretical information, yes. However, writing question items for an imaginary group of learners cannot provide enough practice, of course. Therefore, it can be better to give pre-service students more opportunities to be included in the assessment process and let them complete their theoretical knowledge with practical knowledge (T2).

Having a similar point of view, one of the interviewee claimed in the following way:

No, it hasn't met my needs. The very first reason of this is that I haven't practised any kinds of these assessment types. They were given on a theoretical basis (T4).

In similar vein, the following comment shows the discontent with the contribution of LTA course to their LAL development reported by in-service EFL teachers:

“Not exactly, it was just an introduction” (T5).

In line with many research studies within relevant literature, the participants self stated that they have had little or no training for LAL. Teachers are generally found having low assessment literacy levels in many research studies (Hasselgreen, Carlsen & Helness, 2004; Tsagari & Vogt, 2017; Xu & Brown, 2016) as assessment courses offered at teacher education programs are criticized as being too theoretical rather practical. Moreover, prominent research studies underline that language assessment courses offered at pre-service teacher education program do not improve learners’ practical skills as they are mostly criticized being theory oriented (Kaya & Mede, 2021).

To reach an overall idea of the LTA courses at pre-service language teacher training programs; the content, teaching resources, weekly topics covered in the LTA syllabus of the participants were examined critically with an aim to reveal the most frequently taught topics, the most frequently used teaching materials and the most frequently used assessment type. In doing so the research study aims at picturing the LAL issue by analyzing the related LTA course which is the main figure for the LAL development of prospective EFL teachers. Within this scope, in an attempt to analyze the LTA course the participant student teachers have taken, course syllabus of that state university was examined through document analysis. The syllabus of the course including content, objectives, assessment procedure, and weekly schedule was reached from the university website.

Generally speaking, the majority of the ELT departments in Türkiye offer a specific language testing course in addition to the compulsory Measurement and Evaluation course (taught in L1) offered at other teacher education programs with exceptions (CoHE, 2006). Not only the name of this specific course varies (e.g. Testing in English Language Teaching, English Language Testing and Evaluation, Testing and Evaluation in Foreign Language Teaching, Test Preparation in English Language Learning, etc.) but also the timing (mostly in the 8th term, rarely 7th and 2nd terms) of it may differ across universities

(Şahin, 2018). Therefore, it can be challenging to examine the syllabi and the content of each pre-service language teacher education program as a result of this variety. But still, as put forward by Şahin (2018), the place, timing, and content of the course are noteworthy suggesting some clues for the preparedness level of prospective teachers, the perceived importance levels of LTA concepts and topics. Although the concerns about the time and name of the course are not the issues handled in this dissertation, still then they might have an impact on the perspectives of the participants.

More specifically, the main concern of this study is about the content of the LTA course to name the topics it covers or it does not cover. Because LAL levels of language teachers have been widely investigated and found moderate which results in further questions about the contribution or impact of LTA courses on prospective teachers at pre service teacher education. As mentioned before, the time and place of the course were questioned by Şahin (2018) suggesting the 8th semester might cause some deficiencies at practicum course as a result of the lack of language assessment related knowledge even though the majority of the instructors (14 out of 21) found the 8th term appropriate time to teach LTA in her thesis study. And further only when asked about the number of the course, 12 of them thought sufficient, however during interviews one LTA course was considered inadequate to cover both theoretical and practical knowledge due to the time limit. Therefore, it can be seen that Şahin (2018) proved a dilemma in her study, that is even though the majority of the participants (more than half) were satisfied with the number of the LTA course, they complained about lack of time to cover practical skills and stated a wish for additional topics they would like to adapt to the curriculum. More importantly, Şahin (2018) drew attention the suggestions offered by the instructors which are worth mentioning as follows: (a) one theoretical course to be offered in 7th term covering certain topics such as techniques of assessing productive skills, (b) a second course offered in 8th term covering different LAL related topics leaving room for practice.

Remarkably, as Giraldo (2018, 2020) highlights; teachers' awareness of LAL should be raised during particularly pre-service language teacher education through language assessment courses for the following reasons:

1. Teachers are responsible for “planning, implementing, and interpreting language assessments”

2. They are the ones to communicate assessment results and make decisions about student learning

3. The general conclusion of many research studies show that language teachers are found to have limited or low levels of LAL and need further training to improve their LAL (Chung & Nam, 2018; Vogt & Tsagari, 2014).

Considering these reasons, the objectives, outcomes, and the content of a language assessment and testing course bear greatest importance. In this sense, the following paragraphs will make a detailed analysis of LTA course (Test Preparation in English Language Teaching) offered at the state university from which the student data was collected. To start, the learning objectives as stated in the syllabus suggest that the pre-service teacher candidates are initially offered an introductory week identifying basic terminology and concepts with regards to LTA. When investigated in detail, the main learning objectives as displayed in the course documents can be noted as understanding the importance and role of testing in foreign language teaching by learning the main principles of language testing with an aim to help learners develop a language test. The course mostly concentrated on enabling prospective language teachers to develop a language test by achieving to master the main principles and techniques of language testing. In this sense, the outcomes are considered as follows: At the end of the course, the teacher candidates will be able to identify basic concepts and terms of assessment such as validity, reliability, washback effect, etc., to be aware of testing and teaching relationship, to explain assessment purposes, to know alternative assessment tools, to assess language skills, to prepare a blueprint, to evaluate a language test, to make item analysis and design a suitable test among others.

Concerning weekly content covered in LTA course, it can be witnessed that the first two weeks are devoted to teach key concepts in testing and the purposes of assessment. When closely scrutinized, it might be seen that approximately during the first 8 weeks of the academic term theoretical knowledge of language assessment (types of assessment,

principles and techniques of assessment, alternative assessment (portfolio, peer and self-assessment, reliability, validity) are attempted to teach while the rest of the weeks are devoted to teach practical aspects such as preparing different question types (short answers questions true/false questions multiple choice questions, matching questions), test preparing techniques and testing different language skills. However, as it is demonstrated in the syllabus and supported by the interview data of student teachers, the LTA course is mostly focused on theoretical knowledge leaving less even no room for actual practice. When it comes to the certain topics covered in LTA course, generally as the academic term allows, 14 weeks of instruction are divided to teach the main concepts and topics related to language testing such as validity, reliability, washback, self and peer assessment, to name a few. While the first week of instructions generally seemed to cover introduction to language testing involving the key concepts, purpose and types of assessment, the following weeks focus on the practical aspect of assessment from the alternative assessment methods to preparation techniques of different types of exams.

Concerning the assessment of different language skills, the pre-service student teachers mentioned they have a very high need for assessing speaking and writing as these skills necessitate a more objective and critical perspective. Accordingly, when the syllabus was investigated in detail, it can be seen that only one week (11th week) is devoted to teach assessment of receptive and productive skills. However as revealed by the findings, both pre-service and in-service EFL teachers need more training in those topics. Therefore, it can be assumed that while designing the LTA curriculum, more than two weeks of instruction need to be allocated to teach how to assess speaking and writing, thus two weeks of instruction can be devoted to teach theoretical understanding of the topics while other two weeks can be used for actual practice. Concerning with the teaching method stated in the course documents, it can be seen that students' presentations, sample test analysis and discussion are the main tools to develop LAL of prospective teachers. For assessment of the course; midterm exam, final project, individual study and attending lectures are written as the major ways for the final grade.

In a synthesis, the course documents -the syllabus- including weekly topics, learning objectives, and outcomes reveal the amount of importance given to each language

testing related topic, the amount of time allotted to practice, the type of assessment which in turn suggests the quality and quantity of LAL training future teachers are equipped with.

As both the quality and quantity of LAL training matters for an effective and sound assessment, the following table summarizes the results of training and LAL levels of in-service and pre-service EFL teachers.

Table 24

A comparison of training and perceived LAL level of participants

Received Training		Perceived Level of LAL					Total
		Poor	Fair	Average	Good	Excellent	
Pre-service EFL Teachers	No training	0	1	1	2	0	4
	A little training	2	3	23	12	4	44
	Extensive training	0	0	12	31	5	48
	Total	2	4	36	45	9	96
In-service EFL Teachers	No training	0	7	1	2	0	10
	A little training	2	28	4	33	0	67
	Extensive training	0	2	0	15	4	21
	Total	2	37	5	50	4	98

As the Table 24 demonstrates above, the majority of in-service EFL teachers ($n=67$) reported having a little training while half of the pre-service EFL teachers ($n=48$) noted that they have had an extensive training in terms of LAL. Concerning the training and the levels of pre-service EFL teachers, one can conclude that the majority of them ($n=23$) who reported having a little training ($n= 44$) regarded their level as average. When extensive training is concerned ($n=48$), it might be seen that the majority of the participants ($n=31$) reported themselves having a good level. In similar vein, concerned with the information about training and LAL levels of in-service EFL teachers, it is observed that one half of the informants who reported having a little training ($n= 67$) regarded their level as good ($n=33$) while the rest of them self stated their level as fair ($n=28$).

When it comes to the general overview of the comparison of training and LAL needs, the following table summarizes the main findings for both in-service and pre-service EFL teachers.

Table 25

A comparison of training and perceived need of participants

Received Training		Perceived Need of Training for LAL					Total
		No Need	Low Need	Moderate Need	High Need	Very High Need	
Pre-service EFL Teachers	No training	0	0	2	1	1	4
	A little training	2	8	21	12	1	44
	Extensive training	0	13	24	8	3	48
	Total	2	21	47	21	5	96
In-service EFL Teachers	No training	0	1	5	4	0	10
	A little training	1	6	37	21	2	67
	Extensive training	0	12	5	2	2	21
	Total	1	19	47	27	4	98

According to the results demonstrated above in Table 25, the majority of participants in both groups (pre-service and in-service EFL teachers) reported themselves in moderate need of training for LAL with the percentages 48.5% and 48.0% respectively. Moderate need is followed by high need in each group with 27 out of 98 participants in in-service teachers group and 21 out of 96 in pre-service teachers group of participants. When their received training is taken into account, the results show that the in-service EFL teachers having received extensive training ($n=21$), self reported to be in low need of training ($n=12$) for LAL development while a little training receivers ($n=67$) in the same group of participants regarded their needs in moderate ($n=37$) or high need category ($n=21$). However, when the pre-service EFL teachers are concerned, the issue is different in that even though they claimed to have extensive training ($n=48$), but still they considered themselves in moderate need ($n=24$), high need ($n=8$) and very high need ($n=3$) of training for their LAL development. This might be explained by the fact that their high self efficacy levels and self confidence in terms of LAL, as a result of the recent theoretical knowledge they have received in the time of data collection which might have an impact

on their responses for the training they have received. Yet still, for the following question with regards to the need of training for LAL, they regarded their needs moderate or high/very high which might be stemmed from their feeling of discomfort about the lack of classroom experience and the gap between theoretical and practical knowledge which is also discussed by Öz and Atay (2017) showing the difference between the perceptions about assessment and assessment practices.

4.6. Results and Discussion for Research Question 5 and Research Question 6

What are the language assessment-related topics that participants recognise as important for their language assessment literacy development? (RQ5)

What are the most challenging topics in terms of language assessment literacy? (RQ6)

The results with regards to the language assessment topics will be analyzed under these two research questions together as both of them address the related concepts and topics to be covered in language assessment training in pre-service teacher education. While discussing the issue, the qualitative data gathered via two interview questions (Of the topics above, which one / ones do you think the most helpful to language teachers? Of the topics above, which one / ones do you think the most challenging for language teachers?) also throws shed light on the way. In this regard, the following table summarizes the perceptions of pre-service EFL teachers involving the percentages for the items and number of participants.

Table 26

Pre-service EFL teachers' perceived importance levels of LAL topics

LAL Items	Unimportant		Not very important		Fairly important		Important		Essential	
	N	%	N	%	N	%	N	%	N	%
History of Language Testing	17	17.5	22	22.7	30	30.9	19	19.6	9	9.3
Procedures in language test design	1	1.0	2	2.1	11	11.3	34	35.1	49	50.5
Deciding what to test			1	1.0	4	4.1	20	20.6	72	74.2
Writing test specifications/blueprints	3	3.1	2	2.1	21	21.6	33	34.0	38	39.2
Writing test tasks and items					9	9.3	26	26.8	62	63.9
Evaluating language tests			4	4.1	2	2.1	17	17.5	74	76.3
Interpreting scores			1	1.0	9	9.3	26	26.8	61	62.9
Test analysis			2	2.1	16	16.5	28	28.9	51	52.6
Selecting tests for your own use	1	1.0			17	17.5	36	37.1	43	44.3
Reliability					7	7.2	15	15.5	75	77.3
Validation			1	1.0	2	2.1	16	16.5	78	80.4
Use of statistics	2	2.1	7	7.2	20	20.6	32	33.0	36	37.1
Rating performance tests (speaking/writing)			2	2.1	14	14.4	30	30.9	51	52.6
Scoring closed-response items			7	7.2	25	25.8	37	38.1	28	28.9
Classroom assessment			2	2.1	15	15.5	32	33.0	48	49.5
Large-scale testing	1	1.0	7	7.2	29	29.9	31	32.0	29	29.9
Standard setting	1	1.0	4	4.1	25	25.8	37	38.1	30	30.9
Preparing learners to take tests			4	4.1	12	12.4	31	32.0	50	51.5
Washback on the classroom			6	6.2	11	11.3	32	33.0	48	49.5
Test administration			2	2.1	18	18.6	37	38.1	40	41.2
Ethical considerations in testing			3	3.1	15	15.5	18	18.6	61	62.9
The uses of tests in society	2	2.1	9	9.3	26	26.8	30	30.9	30	30.9
Principles of educational measurement	1	1.0	4	4.1	10	10.3	34	35.1	48	49.5

As displayed in Table 26, the majority of pre-service EFL teachers (80.4%) thought that “validation” was the most required topic to cover in LTA course during pre-service teacher education, while perceiving “history of language testing” as less required (17.5%). When the results are analyzed closely, it can be seen that “reliability”, “evaluating language tests”, and “deciding what to test” are the next topics which have the highest percentages as 77.3 %, 76.3 % and 74.2% respectively based on the self responses of

informants. On the other hand, “use of statistics” and “large scale testing” items were among the least important topics which were stated by pre-service EFL teachers.

Considering the in-service EFL teachers’ opinions with regards to the importance level of LAL topics, the following table is prepared in order to demonstrate the main findings.

Table 27

In-service EFL teachers’ perceived importance levels of LAL topics

LAL Items	Unimportant		Not very important		Fairly important		Important		Essential	
	N	%	N	%	N	%	N	%	N	%
History of Language Testing	15	15.3	18	18.4	35	35.7	14	14.3	16	16.3
Procedures in language test design	4	4.1	11	11.2	34	34.7	49	50.0	4	4.1
Deciding what to test					6	6.1	26	26.5	66	67.3
Writing test specifications/blueprints					4	4.1	39	39.8	55	56.1
Writing test tasks and items					5	5.1	29	29.6	64	65.3
Evaluating language tests					5	5.1	28	28.6	65	66.3
Interpreting scores					4	4.1	25	25.5	69	70.4
Test analysis			1	1.0	7	7.1	38	38.8	52	53.1
Selecting tests for your own use			3	3.1	9	9.2	32	32.7	54	55.1
Reliability			2	2.0	3	3.1	22	22.4	71	72.4
Validation			2	2.0	4	4.1	28	28.6	64	65.3
Use of statistics			8	8.2	23	23.5	29	29.6	38	38.8
Rating performance tests (speaking/writing)			3	3.1	6	6.1	26	26.5	63	64.3
Scoring closed-response items	4	4.1	5	5.1	13	13.3	36	36.7	40	40.8
Classroom assessment	1	1.0	2	2.0	9	9.2	33	33.7	53	54.1
Large-scale testing	1	1.0	5	5.1	24	24.5	39	39.8	29	29.6
Standard setting	1	1.0	2	2.0	18	18.4	38	38.8	39	39.8
Preparing learners to take tests	1	1.0	2	2.0	16	16.3	31	31.6	48	49.0
Washback on the classroom	1	1.0	4	4.1	18	18.4	28	28.6	47	48.0
Test administration			2	2.0	13	13.3	35	35.7	48	49.0
Ethical considerations in testing	1	1.0	2	2.0	5	5.1	29	29.6	61	61.2
The uses of tests in society	2	2.0	6	6.1	24	24.5	37	37.8	29	29.6
Principles of educational measurement			3	3.1	15	15.3	32	32.7	48	49.0

The majority of in-service EFL teachers thought that “reliability” was the most required topic to cover for LAL development, while perceiving “history of language testing” as less required. According to the results summarized above in Table 27, the 72.4% of in-service EFL teachers thought that “reliability” was the most important topic to build LAL knowledge. When the results are analyzed closely, it can be seen that “reliability” is followed by “interpreting scores” and “deciding what to test” which have the highest percentages as 70.4 % ($M=4.66$, $SD=.555$) and 67.3 % ($M=4.61$, $SD=.603$) respectively based on the self responses of informants. On the other hand, “history of language testing” and “the uses of tests in society” were among the least important topics which were stated by in-service EFL teachers with the mean values $M=2.98$ ($SD=1.268$) and $M=3.87$ ($SD=.981$) respectively.

Consideration of the qualitative data in terms of the challenging and useful LAL topics discussed under the 6th research question, the following table is prepared to display the results of pre-service EFL teachers’ perceptions.

Table 28

The perceptions of pre-service EFL teachers about LAL topics

Research Foci	Related LAL topics	N	%	Interviewee Codes
The most challenging LAL topics	Reliability	2	100%	S1, S2
	Validation	2	100%	S1, S2
	Rating performance tests (speaking/ writing)	2	100%	S1, S2
	Interpreting scores	2	100%	S1, S2
	Validation	2	100%	S1, S2
The most useful LAL topics	Reliability	2	100%	S1, S2
	Writing test specifications/ blueprints	1	50%	S1
	History of language testing	1	50%	S1

In pursuit of exploring the most useful and challenging LAL topics self reported by pre-service EFL teachers, content analysis was done and the recurrent themes were drawn from the raw data which were summarized and exemplified above in Table 28. According to these results, reliability and validation seem outstanding as both in terms of challenging and usefulness aspects. On the other hand, the last two LAL items in both lists differ among interviewees.

Initially, the general tendency portrays that rating performance tests and interpreting scores were perceived as challenging which revealed that pre-service EFL teachers feel unprepared and unable enough to administer such productive tests and reach a clear understanding of the results to produce meaningful feedback. When analyzed closely, the following comment implies the underlying reasons for this perspective:

In open ended questions within these productive tests, it will be difficult to rate and interpret them as a result of variety of answers (S1).

In similar vein, an outstanding comment from S2 in association to RQ8 is actually widening horizon to the challenges pointed out by S1. To exemplify, S2 highlighted that;

There should be rubric which clarifies what to do, how to do, and the way to be assessed. (...) Then, there will be no question as the rubric guides the learner and teacher as well (S2).

More saliently, the items noted as useful also should be paid attention to throw shed light on the underlying beliefs of pre-service EFL teachers. Considering “history of language testing”, the interviewee -S1- acknowledged that being aware of the evolving stages of assessment methods enables teachers to carry out a well-designed test which is explained in his own words as follows:

We can see the pros and cons of a language assessment method by learning its evolving stages (S1).

With regards to the qualitative data in terms of the challenging and useful LAL topics discussed under the 6th research question, the following table is prepared to demonstrate the results of in-service EFL teachers' perceptions.

Table 29

The perceptions of in-service EFL teachers about LAL topics

Research Foci	Related LAL topics	N	%	Interviewee Codes
The most challenging LAL topics	Use of statistics	6	42.8%	T3,T4,T5,T6,T8,T10
	Writing test tasks and items	6	42.8%	T2,T3,T5,T8,T10, T14
	Rating performance tests (speaking/writing)	5	35.7%	T2,T5,T7, T12, T14
	Interpreting the scores	5	35.7%	T2,T3,T7, T8,T9
	Classroom assessment	3	21.4%	T2,T6,T9
The most useful LAL topics	Validation	8	57.1%	T2,T3,T4,T5,T7,T8,T9,T14
	Reliability	7	50%	T2,T3,T4,T5,T7,T9, T14
	Deciding what to test	7	50%	T3,T6,T7,T9,T10, T12, T14
	Writing test tasks and items	5	35.7%	T3,T4,T5,T7,T9
	Interpreting scores	5	35.7%	T3,T7,T9,T10, T13

Having identified how important the LAL related topics are according to the responses of in-service EFL teachers, now it is of utmost importance to determine the most challenging and the most useful LAL topics for EFL teachers. In doing so, it is quite possible to decide on the content of LTA course and arrange the topics accordingly. The general tendency of interviewees' thoughts about the issue showed that "Classroom assessment", "Use of statistics", "Writing test tasks and items", "Interpreting the scores",

and “Rating performance tests (speaking/writing)” were found as the most challenging topics. The highest percentages belong to the “use of statistics” (42.8%) and “Writing test tasks and items” (42.8%) which imply that the participant teachers generally have difficulty in applying statistics during language assessment process. Moreover, productive skills are always mentioned the most difficult skills to assess because of the fact that assessing these skills requires objective evaluation without rater’s bias and specific content knowledge. This general tendency was also witnessed in the findings as the 35.7% of the in-service teachers regarded rating productive skills as challenging.

Concerning the most useful LAL related topics, even though the interviewees had a different list of items, the most mentioned five topics can be listed in the following way: reliability, validity, writing test tasks and items, deciding what to test and interpreting scores. The analysis of each item points out that the highest percentages were found with regards to “validity”, “reliability” and “deciding what to test” as 57.1%, 50% and 50% respectively. In pursuit of interpreting these findings, it can be concluded that the items found as the most useful are generally related to development of an assessment task and/or tool which requires to be valid and reliable. That’s why the EFL teachers mostly self reported the importance of having necessary knowledge and skills associated with ensuring reliability and validity of a test. What’s more, “writing test tasks and items” and “interpreting scores” are found very useful by the 35.7 % of the interviewees. Two more participants also stated “rating performance tests” (T12, T14) and “ethical considerations in language testing” (T12, T13) as helpful for their LAL development and a sound language assessment.

Certain language assessment topics and items have been found challenging to acquire and apply in various studies indicated in related literature as follows: validity, reliability, interpreting assessment results (DeLuca & Klinger, 2010; Mertler, 2009), using different assessment types (Volante & Fazio, 2007) and providing constructive feedback (Giraldo, 2019). As a matter of fact that, the points shared in Mertler’s study (2009) incorporate the findings of this study in that interpreting assessment results were regarded as challenging by the in-service EFL teachers within this current study. Moreover, assessing productive skills was also found challenging while providing feedback was not regarded as problematic in this study.

4.7. Results and Discussion for Research Question 7

What are the possible ways for language assessment literacy development? (RQ7)

Considering this research question, participants were asked initially “What does LAL mean for you? (IQ10)”, and then “How do you improve LAL? (IQ4),” and “What are the possible ways for improving language assessment literacy? (IQ5)” questions were posed during the interview.

Considering the meaning of LAL, the 10th interview question tried to explore the perceptions of each group of participants which were summarized below in Table 30.

Table 30

The main findings for IQ10

Group of Participants	Meaning of LAL	N	%	Interviewee Codes
Pre service EFL teachers	Knowledge-based meaning	2	100%	S1, S2
	Knowledge-Based Meaning	8	57.1%	T2, T3, T4, T6, T8, T9, T11, T13
In service EFL teachers	Skills-Based Meaning	4	28.5%	T1, T2, T4, T8
	Awareness-Based Meaning	2	14.2%	T5, T7

When the qualitative data was coded, it was observed that each interviewee had a sort of meaning in their minds to some extent which were coded as knowledge-based meaning, skills-based meaning and awareness-based meaning. To illustrate, the first theme was voiced by the majority of the participants (10 out of 16) some of which were exemplified below:

It is the mastery of knowledge and skills in designing and developing assessment tasks and using the data obtained via assessment (T4).

It is what teachers know, should know and do not know about language assessment (T3).

As a more comprehensive identification, Teacher 5 proposed the following definition in her own words:

Being aware of (...) the procedures in assessment process, what methods can be used to assess learners, what ethical issues should be considered, how to develop, adapt or adopt assessment materials, reliability and validity (T5).

As the interviewees consider LAL from knowledge, skill and awareness based domains; the related literature is also in the same vein as portrayed in the following definition put forward by Coombe, Vafadar, and Mohebbi (2020: 2):

Language assessment literacy is generally viewed as a repertoire of competences, knowledge of using assessment methods, and applying suitable tools in an appropriate time that enables an individual to understand, assess, construct language tests, and analyze test data.

Having identified the overall meaning LAL has according to the participants, now it is time to learn how they develop their LAL levels. In an attempt to explore the main ways for developing LAL, the following table is designed based on both groups of participants' responses.

Table 31

Main themes identified for IQ4 and IQ5 in “needs for assessment” category

Group of Participants	Ways for LAL development	N	%	Interviewee Codes
Pre-service EFL teachers	Reading relevant literature	2	100%	S1, S2
	Academic meetings (Seminars, workshops, conferences, webinars, etc.)	2	100%	S1, S2
	Professional development programs (In-service trainings etc.)	1	50%	S1
	Self-reflection and personal experience	1	50%	S2
In-service EFL teachers	Academic meetings (Seminars, workshops, conferences, webinars, etc.)	6	42.8%	T1, T3, T4, T5, T10, T11
	Reading relevant literature	6	42.8%	T2, T4, T5, T6, T7, T9
	Professional development programs (CELTA, DELTA, in service trainings etc.)	5	35.7%	T3, T7, T8, T9, T10
	Graduate Education (Master, PhD)	5	35.7%	T1, T3, T5, T8, T11
	Self-reflection and personal experience	1	7.1%	T12

When analyzed qualitatively, the data indicates that in-service EFL teachers mostly self reported the following ways: academic activities (such as seminar, conference, and workshop), reading relevant literature, professional development programs, and further graduate education (such as Master and PhD programs) with the following percentages 42.8%, 35.7%, 65% and 44% respectively. As displayed in Table 31 above, pre-service EFL teachers’ data corroborates these findings. Thus, academic meetings and professional development programs can be very beneficial for teachers’ LAL awareness and development. However, as teachers play an active role through engaging and collaborative activities, continuous professional development programs contribute a lot compared to one shot workshop as stated by Wiliam and Thompson (2008). In similar vein, Saputra, Hamied, and Suherdi (2020) examined the contribution of a professional learning community on teachers’ language assessment literacy reporting a positive change in teachers’ beliefs about LAL and enhancing the use of authentic assessment. Based on the self reports of participants, this professional program helped them improve the

understanding of assessment ‘for’ and ‘as’ learning replacing assessment of learning. Additionally, in another study, Koh et al. (2018) investigated the impact of a two-year authentic assessment professional development program on teachers’ task design in terms of assessment literacy aspect concluding that the program has a positive effect to some extent.

Moreover, as one of the interviewees emphasized peer feedback, collaboration among colleagues (termed as collegiality) also pave way to LAL development which is also supported by Babaii and Asadnia (2019) in their research study which used asynchronous online discussions, reflective narratives, semi-structured interviews, and scenarios. Thus, reflective thinking and collaborative activities are supported as they foster language assessment of teachers through professional development programs which was exemplified by the following excerpt:

While pre-service training can help teacher candidates learn about basic of assessment, online resources, in-service training and postgraduate studies can help teachers deal with everchanging demands of assessment (T8).

Moving the discussion further, from a socio cultural point of view, some studies highlighted “contextualized, collaborative, and reflective” nature of LAL. For example, some significant figures supported contextual, experiential concerns (Yan & Fan 2020), self-reflection and apprenticeship (Scarino, 2013), collaborative and reflective assessment practices (Harding & Kremmel, 2016; Levi & Inbar-Lourie, 2020; Vogt & Tsagari, 2014) for LAL development. Yan and Fan (2020) pointed out that language teachers can improve their LAL through collaborative assessment activities and communication with language testing researchers, assessment coordinators, teacher mentors, or experienced instructors in an apprenticeship-based, experience-mediated model. As previously noted, both contextual and experiential factors shape assessment practices of language teachers by influencing their tendency to use a certain type or tool for assessment as they mostly prefer to use familiar assessment methods they have experienced as test takers (Reynolds-Keefers, 2010). Thus, contextual aspect of assessment represents a wide range of

parameters from educational to historical and social life while experiential aspect, as the name suggests, is related with the individual's assessment background involving training, real life practices, beliefs and one's own experiences of assessment.

Based on the respondents' views with respect to LAL development and the brief discussion above, the following figure was designed by the researcher to draw a main framework to describe the key considerations in LAL development process which is believed to contribute to the design of LAL training programs.

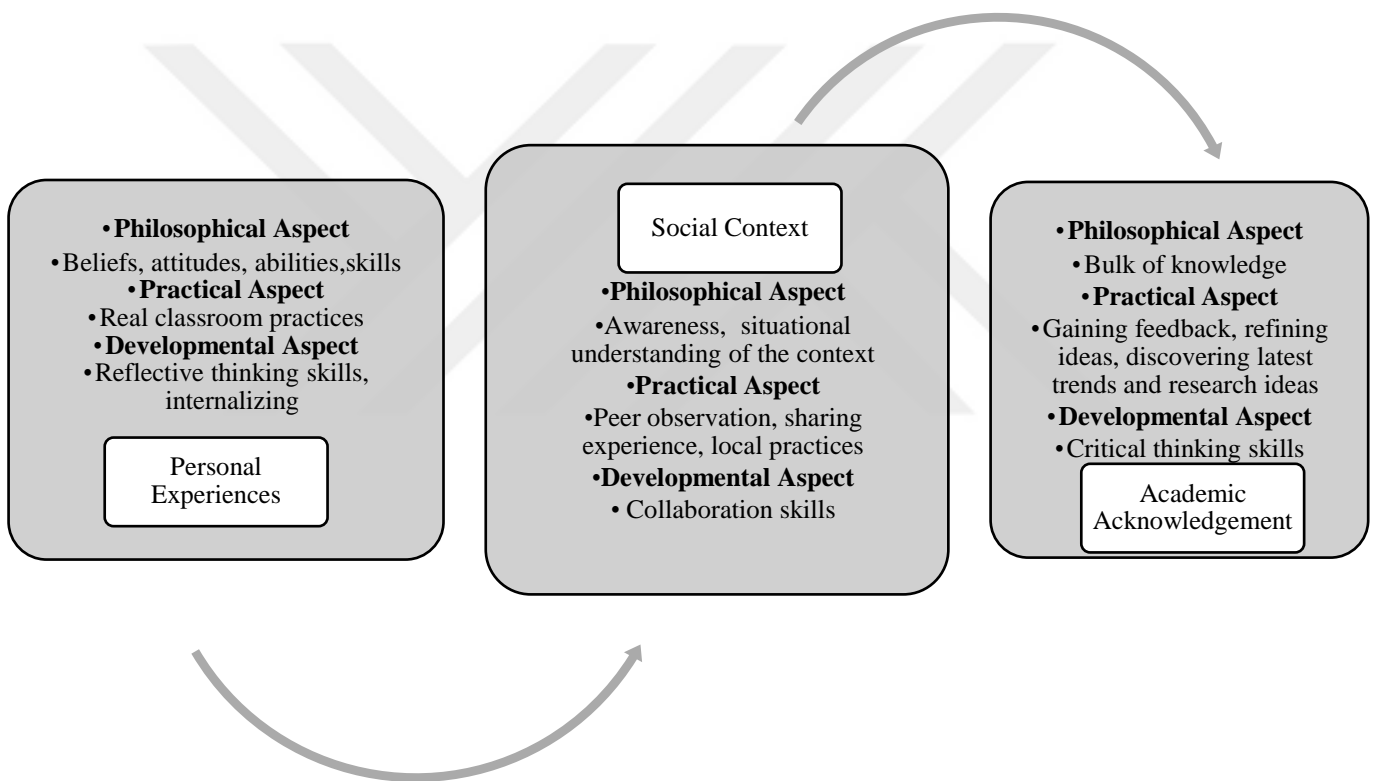


Figure 4. A framework for language assessment literacy development

In order to enhance the efficacy of training for LAL development, the Figure 4 was designed by the researcher to handle the issue from philosophical, practical and developmental aspects. In doing so, teachers' personal experiences, social context and academic acknowledgement, which are believed to shape the language assessment practices, will be discussed through these three abovementioned domains.

To start, the first domain named as “personal experiences” bears greatest importance thanks to its being core entity for a language assessment practice. That’s to say, from philosophical aspect, one’s beliefs, attitudes, abilities and skills shape the way they design a test, the way they apply it, the way they interpret and communicate the results to test-takers. For this reason, while designing a LAL development training program, all stakeholders need to take into consideration philosophical aspect of individuals. Concerning the second domain, which is called practical aspect; personal experiences place a strong impact on classroom practices in terms of language assessment. Lastly, developmental stage implies the necessary qualities for a sound and meaningful language assessment practice. In this case, the first feature in the figure -personal experiences- necessitates reflective thinking skills and internalizing process for an individual to improve her/ his LAL. On the contrary to what theory based teaching offers, practical knowledge provides opportunities for hands-on experiences to bridge theory-practice gap as active implementers of assessment requires to know how to put into practice the assessment related theories. To put it short, in order to benefit from personal experiences with an attempt to enhance LAL, one needs to reflect on his/her language assessment practices identifying weakest and problematic areas so as to understand the assessment process thoroughly.

When it comes to the second essential part in LAL development, named as social context, it is quite necessary to comprehend the importance of the context which embodies the setting, the people, the language policies, and etc. Therefore, for the first stage-philosophical aspect-, being aware of the abovementioned issues (setting, people, policy, etc.) and situational understanding of the context are the main ingredients for a successful LAL development. To exemplify, within the philosophical stage of social context domain; test purpose of a language assessment task can be achieved successfully when test developer is aware of the target group, individual differences of test takers, the language policy embedded there, among others. Additionally, for practical stage of social context, some certain activities such as peer observation or sharing experience could be embraced in order to improve language assessment practices. Moreover, the last stage (developmental aspect) requires collaboration skills to make use of peer interactions and experiences. To illustrate, as proposed by Yan and Fan (2020), self-reflection and

apprenticeship also might facilitate LAL development in addition to contextual and experiential tenets. More comprehensively, “mentors, colleagues and published assessment materials” (Tzagari & Vogt, 2017: 52), “exchanging knowledge and experience, observing others’ assessment practices, asking for colleagues’ guidance” (Babaii & Asadnia, 2019: 757), reflections and assessment practices (Levi&Inbar-Lourie, 2020) can be regarded beneficial ways for LAL development which were also mentioned by the participants in this current study.

Last but not the least, the third outstanding pillar, “Academic Acknowledgement”, was given a place as the first two domains (personal experiences and social context) build a bulk of knowledge ultimately. In order to build such knowledge, one needs to gain feedback, refine ideas, discover latest trends and research topics at the practical stage and further develops LAL skills and competences through critical thinking skills at the developmental stage. When the qualitative data examined in detail, it might be seen that almost half of the participants (42.8 %) regarded academic knowledge (conferences, seminars, post graduate studies, and reading articles, to name a few) as the most beneficial for their LAL development.

Considering the implementation of this framework in practice, what is important to recognise is that every one of the stages within each domain fulfils different functions. To illustrate, philosophical stage takes a proactive approach in which familiarizing with the related phenomenon is attempted (e.g. familiarizing with beliefs in ‘Personal Experience’ or familiarizing with the setting, local policies in ‘Social Context’). With regards to ‘Practical stage’, it can be considered as a stage where the actual use of the items in philosophical stage within the aforementioned domains or the stage where the relevant items might be witnessed and utilized. For example, having familiarized with the local setting in the first stage (Philosophical), peer observation in practical stage can be used as a strategy or technique in order to benefit considerably from social context in pursuit of developing LAL. Finally, developmental stage is generally related to the necessary qualities which can be used to create certain activities or tasks like critical thinking activities, collaborative activities or reflective activities to enhance LAL. For example, reflective journals can be utilized for a sound and meaningful language assessment practice

in real classroom settings by promoting reflective thinking, so that one can understand the weaknesses and strengths of his / her assessment practice.

In short, the framework designed in this current study can be considered as one the major contributions of the present study as it provides a comprehensive guide to develop LAL from philosophical, practical and developmental aspects.

4.8. Results and Discussion for Research Question 8

What are the pre-service and in-service language teachers' perceptions about online language assessment? (RQ8)

For RQ8, five interview questions will be discussed together as they deal with online language assessment (OLA, henceforth) in general sense including online language assessment tools, tasks, online feedback, etc. Within this regard, it was intended that interview questions would provide information connected to the aforementioned subjects based on the comments of interviewees.

Therefore, the following table attempts to summarise and exemplify the main findings based on pre-service EFL teachers' responses including research foci, main themes and categories as well which were arranged below.

Table 32

Main themes identified for “online language assessment” by pre-service EFL teachers

Research Foci	Category	Main Themes Emerged	N	%	Interviewee Codes
Opinions of OLA	Disadvantages	Academic misconduct	2	100%	S1, S2
		Difficulty of proctoring	2	100%	S1, S2
		Reliability-validity concerns	2	100%	S1, S2
	Advantages	Stress-free	1	50%	S1
Suitable Tasks for OLA	Higher-order thinking tasks	Individual projects	2	100%	S1, S2
		Research based tasks	2	100%	S1, S2
	Interactive tasks	Role-plays	1	50%	S1
		Group works	1	50%	S2
Challenges in OLA	Student-related	Academic misconduct	2	100%	S1, S2
		Lack of technological knowledge	1	50%	S2
		Distraction	1	50%	S1
	Teacher related	Lack of technological knowledge	1	50%	S2
	Technology related	Internet connection	2	100%	S1, S2
		Lack of devices	2	100%	S1, S2
Coping Strategies	Technology-aided	Proctoring programs (Turnitin,etc.)	2	100%	S1, S2
	Teacher-oriented	Use of cameras	2	100%	S1, S2
		Rubric development	1	50%	S2
Online feedback	Disadvantages	Types of questions and tasks	2	100%	S1, S2
		Lack of communication	2	100%	S1, S2
	No difference	Feedback provider	1	50%	S1

To discuss “online language assessment” from all possible angles, five interview questions were formulated to elicit responses on the following issues as “overall idea for

OLA, suitable tasks and tests for OLA, online feedback, challenges in OLA and coping strategies if any". To fulfil the purposes above, the responses obtained through interviews were coded and the common themes were grouped under related category. In this regard, as displayed in Table 32 above, the data linked to the first category -overall idea of OLA-, both positive and negative perspectives were mentioned with various themes and focus points. For instance, negative statements were related to the difficulty of academic misconduct and proctoring problems along with reliability and validity concerns in OLA practices while positive feature was regarded as stress-free nature of OLA.

Considering suitable tasks and tests for OLA, the findings revealed two categories as "Higher-order thinking tasks" and "Interactive tasks". In each category, certain language assessment tools were voiced by the pre-service interviewees. For the first category, "Individual projects", "Research based tasks" and "Portfolios" were considered suitable and favoured by the interviewees. When interactive tasks are concerned, role plays and group works in online platforms are the mostly mentioned tasks as they were believed to enable learners to produce the language within a context.

When it comes to the challenges in OLA, the findings of IQ1 and IQ4 support each other. To start, the qualitative data obtained through interviews with regards to the challenges are categorized as student-related, teacher-related and technology-related challenges. In relation to student-related problems, it can be seen that academic misconduct stood out as a model response in addition to lack of technological knowledge and distraction. Information and communication technology (ICT, henceforth) related knowledge is believed to facilitate the online teaching and learning as well as online language assessment according to the interviewees which can be exemplified as follows:

OLA became ineffective and inferior due to the lack of ICT knowledge of both students and teachers in emergency technology integrated learning (online education) which would be an opportunity instead (S2).

Moreover, the evidence collected also uncovered that technical problems like internet connection or lack of necessary devices were other challenges both teachers and learners have faced during this experience. To overcome these problems, certain coping strategies were offered by the pre-service EFL teachers as follows: proctoring programs (Turnitin, etc.), use of cameras, rubric development, types of questions and tasks. As an eye-catching comment, it is worthwhile to mention the importance of rubric in an online language teaching and assessment which is uttered as;

There should be rubric which clarifies what to do, how to do, and the way to be assessed. Because of the fact that there is no face to face communication in OLA, rubric serves as a means of communication between teacher and learner. Then, there will be no question as the rubric guides the learner and teacher as well (S2).

When online feedback is considered, generally speaking the pre-service interviewees have shared negative concerns. To exemplify, one of the interviewees (S1) mentioned a lack of communication as the main problem for online feedback in his own sentences:

In peer feedback, I asked some questions on my pair's written paper, however I could not get an answer (S1).

Moreover, another interviewee emphasized the importance of feedback provider which is the determining factor for the quality of feedback no matter what way it is provided, either face to face or in an online platform.

With respect to in-service EFL teachers' perceptions about OLA, the following Table 33 displays the main findings.

Table 33

Main themes identified for “online language assessment” by in-service EFL teachers

Research Foci	Category	Main Themes Emerged	N	%	Interviewee Codes
Opinions of OLA	Overall tendency	Being Inevitable	14	100%	T1-T14
	Disadvantages	Academic misconduct	6	42.8%	T2, T4,T5, T7,T11,T12
		Difficulty of proctoring	4	28.5%	T2, T4,T5, T7
		Reliability-validity concerns	3	21.4%	T2, T4, T10
	Advantages	Off-place learning	4	28.5%	T2, T8, T9, T11
		Practical	3	21.4%	T8, T11, T12
		Fun	1	7.1%	T2
Suitable Tasks for OLA	Higher-order thinking tasks	Individual projects	5	35.7%	T3, T5, T9, T10, T11,
		Research based tasks	4	28.5%	T3, T9,T10, T11
		Portfolios	3	21.4%	T5, T9, T10
	Interactive tasks	Dialogues with native speakers	1	7.1%	T1
	Traditional tasks	Multiple choice tests	1	7.1%	T4
		Writing	1	7.1%	T7
	Challenges in OLA	Student-related	Academic misconduct	13	92.8%
Lack of technological knowledge			2	14.2%	T2, T12
Teacher related		Lack of technological knowledge	3	21.4%	T10, T11, T13,
		Internet connection	7	50%	T3, T5, T6, T8, T10,
Technology related		Lack of devices	1	7.1%	T8
		Proctoring programs (Turnitin, Witwiser, etc.)	4	28.5%	T2, T7, T9, T11
		Limited time for each question	1	7.1%	T10
Coping Strategies	Teacher-oriented	Distribution of questions	2	14.2%	T4, T7,
		Types of questions and tasks	3	21.4%	T7, T10, T11
	Advantages	Practical	6	42.8%	T2, T4, T8,T10,T11,T12
Online feedback	Disadvantages	Enjoyable	2	14.2%	T4, T8
		Difficult	5	35.7%	T3, T5, T7, T10, T12
		Time-consuming	2	14.2%	T5, T7

The details presented in this table, compiled from interview data provide patterns of common themes concerned with interviewee opinions with respect to OLA, suitable tasks for OLA, challenges and coping strategies during OLA, and further online feedback. Consideration of the first category as “overall idea of OLA”, both positive and negative perspectives were mentioned with differing themes and focus points. For instance, negative statements were related to the difficulty of academic misconduct (42.8%) and proctoring problems (28.5%) in addition to reliability and validity quality (21.4%) while positive aspect was considered as practical (28.5%) and its being fun (7.1%) and engaging for learners in addition to its off-place nature. The positive sides of online assessment were also emphasized by Elzainy et al. (2020) as enhancing students’ autonomy and their critical thinking skills.

When it comes to suitable tasks and tests for OLA, the findings are categorized in three groups by the researcher as “Higher-order thinking tasks”, “Interactive tasks” and “Traditional tasks”. In each category, certain language assessment tools are voiced by the interviewees. For the first category, favoured by the majority of the participants “Individual projects” (35.7%), “Research based tasks” (28.5%) and “Portfolios” (21.4%) were found suitable and justified by the interviewee (T2) as follows:

“As for the concern about cheating, I would personally choose higher-order thinking questions as they cannot cheat, there is no right or wrong answer and to be able to answer this type of questions, they need to master the topics they are thought” (T2).

As it can be witnessed in the excerpt above, higher order thinking tasks could be more beneficial for online language assessment as they require critical thinking and research skills through individual participation. When interactive tasks are concerned, T1 explained it in her own words: “Having dialogues, like friendship, if someone is abroad (native speaker maybe) etc. being face to face with cameras”. However, even though one of the interviewees (Teacher 7) criticized using multiple choice questions in online language assessment as it enables learners cheat with ease; yet still one another participant

(Teacher 4) supported to use multiple choice questions as they are easy to administer and score.

As obviously witnessed, the challenges during online language education during COVID-19 were mostly observed in assessment process. Therefore, the opinions of interviewees need to be highlighted in order to determine the most challenging issues for online language assessment along with their coping strategies. The qualitative data emerged from the interviews with regards to the challenges are grouped as student-related, teacher-related and technology-related challenges. When the findings for student-related problems were analyzed closely, it can be concluded that the highest percentage is related to “academic misconduct” with a percentage of 92.8% which has been acknowledged by many researchers in related literature. However, Kruger (2015) claimed that online assessment methods can be monitored more easily through some web applications than traditional assessments.

Another important challenge was considered as lack of technological knowledge of both students and teachers voiced by interviewees as ICT related knowledge or technology literacy is believed to have a facilitative impact on online learning and assessment. Moreover, technical problems like internet connection or lack of necessary devices are also among the challenges both teachers and learners have faced during this experience. To overcome these problems, teachers had to find ways to cope with the concerns aroused during online language assessment. These coping strategies can be summarized as follows based on the responses gathered by interviews: proctoring programs (Turnitin, Witwiser, etc.), limited time for each question, distribution of questions, and types of questions and tasks with percentages 28.5%, 7.1%, 14.2% and 21.4% respectively. The challenges and coping strategies seem in alignment with the related research studies. To exemplify, in order to cope with these difficulties, some solutions are offered by Elzainy, El Sadik and Abdulmonem (2020) as follows: time limit for each question and the whole test, designing brief exams, preparing scenario-based questions, randomising questions and answer choices, not permitting to go back to the former question among others.

Contrary to the traditional way of feedback (face to face, written feedback), online feedback was considered by interviewees as advantageous as it was believed to be “practical” and “enjoyable” in general sense. As online feedback provides one to one interaction between learner and teacher either online or via video, it ensures special time for each student. More comprehensively, the advantages of online feedback can be summarized in the following comment of one of the interviewees:

“You can give feedback in video or audio recording format as well as written or coded- feedback options. It is also easier to document and store the feedback you have given” (T8).

It is also advantageous as it provides timely feedback through an online test platform allowing students see their results afterwards.

However, some interviewees voiced concerns about the difficulties and problems with respect to online feedback as exemplified below:

“I believe it is more difficult because giving real time feedback is easier if you and the students are in class, however, for crowded classes it is difficult to give real time online feedback” (T7).

To put them all together, because of the fact that designing online assessment tasks and tests poses a big challenge for teachers, Rahim’s study (2020: 59) can be considered a guiding source for its identifying nine guidelines as in the following:

(a) Evaluate prerequisites for implementing online assessment; (b) ensure alignment of assessment activities with stated learning objectives; (c) address the diversity of students’ situations; (d) maintain a good balance of formative and summative assessments; (e) stimulate student learning with online assessment; (f) consider

format, scheduling and timing of tests; (g) establish clear communication to students regarding assessment matters; (h) ensure high-quality feedback; (i) address assessment validity threats.

4.9. Chapter Summary

As an overview, Chapter 4 presented the results of both qualitative and quantitative data answering each research question. Further, discussing the previous research with an attempt to reach a general understanding of the issue was given a big place within this chapter by reporting both contradicting results and similar findings. Now, it is necessary to put all previous parts together in the last chapter in order to display an overall picture of the issue investigated within this study by drawing attention to significant conclusions.

CHAPTER FIVE

CONCLUDING REMARKS AND IMPLICATIONS

5.1. Chapter Introduction

Within this chapter, an overview of the study with its purpose, methodology, and results is provided. Then, the chapter moves on summarizing the most salient results along with conclusions. Finally, the implications for different stakeholders and future research recommendations will be shared at the end of the chapter.

5.2. Summary of the Study

5.2.1. Purpose of the Study

The purpose of this study is to explore the LAL levels and needs of EFL teacher candidates and in-service EFL teachers, further to question their language assessment beliefs and LAL development. Moreover, pre-service language teacher education programs will be under scrutiny with regards to their contribution to teachers' language assessment literacy development. Considering these basic purposes, making a contribution to the field of LTA by providing valuable information for the stakeholders including program developers, policy makers, teacher educators, and teachers in designing an effective and practical LTA course to improve teacher candidates' language assessment literacy knowledge, skills and competences by covering necessary topics within this course was aimed. With an attempt to realize these purposes, the following research questions were generated:

RQ 1: What are the language assessment literacy needs and levels of pre-service language teachers?

RQ 2: What are the language assessment literacy needs and levels of in-service language teachers?

RQ3: What are the pre-service and in-service language teachers' beliefs about language assessment?

RQ4: What is the contribution of language testing and assessment (LTA) course offered in pre-service language teacher education programs?

RQ5: What are the language assessment-related topics that participants recognise as important for their language assessment literacy development?

RQ6: What are the most challenging concepts of language assessment literacy topics?

RQ7: What are the possible ways for language assessment literacy development?

RQ8: What are the pre-service and in-service language teachers' perceptions about online language assessment?

5.2.2. Summary of the Methodology

As the main purpose of the study was to explore language assessment literacy levels, needs and beliefs, purposive and convenience sampling was used. Therefore, the target groups of participants consist of two main informants as pre-service EFL teachers and in-service EFL instructors working at universities in Türkiye. To this end, the present study is designed as a convergent parallel mixed-methods design (Creswell, 2015) to shed light on the language assessment literacy levels, needs and beliefs of participants from a post positivist perspective thus aiming at unveiling participants' own realities in their naturalistic contexts (Cohen, et al. 2007) thus letting their voices be heard (Inbar-Lourie, 2017).

For this reason, both qualitative data (obtained from interview notes/ recordings and course syllabus) and quantitative data (gathered through a questionnaire) were collected. The qualitative data was analyzed through content analysis using inductive and deductive approaches by grouping codes determining the most common themes (Charmaz, 2014; Weber, 1990). The common emergent themes and topics were identified and grouped in major categories in accordance with the interview questions (Creswell & Poth, 2018; Dornyei, 2007), thus the emergent data align well with the guiding questions within the interviews. For quantitative data, SPSS (version 20) was used to measure the frequencies and percentages for survey questions investigating the LAL levels, LAL needs and perceived importance levels of assessment related topics based on respondents' answers.

5.2.3. Summary of the Results

Research Question 1: What are the language assessment literacy needs and levels of pre-service language teachers?

The first research question invited informants to rate their LAL level with regard to language assessment topics to which 37.1 % of pre-service EFL teachers considered their level as “average” while 46.4 % of them identified themselves as having good level. More specifically, on the matter of “History of language testing”, 58.8 % of informants stated that they feel unprepared and unknowledgeable in terms of the developmental stages of language assessment and testing which might cause the selection of an inappropriate assessment method or task contradicting the purpose of assessment. For instance, one individual without necessary knowledge about formative assessment including the purposes, ways, suitable tasks, etc. might design a language test which actually reflects the summative way of assessment. On the other hand, the majority of the pre-service EFL teachers within the study (84 out of 96) perceived their level for “Deciding what to test” as either good (48.5%) or excellent (38.1%) which might be explained by the fact that the course content has had a positive impact on their knowledge as one of the outcomes of the course was stated as “be able to describe why, when, what to evaluate” in the course documents. As seen clearly, the objective of the LTA course was to equip the future teachers with necessary knowledge about the purpose, content and the timing of the test

which was proved in this paper that the course achieved its goal. This result agrees with DeLuca and Klinger (2010) who also supported the positive impact of assessment course on pre-service teacher assessment literacy development.

In connection with LAL needs, the first research question revealed that 45.8% (44) of pre-service EFL teachers rated themselves as having little training, thus being in moderate or high need category in terms of LAL related topics. Remarkably, “Moderate Need” represented the modal response with 48.5 % ($n=47$) of informants. In total, 70.1 % ($n=68$) of pre-service EFL teachers either perceived their need as “moderate” (48.5%) or “high need” (21.6%). Thus, only 21.6% (21) of participants were left in the ‘low need’ category. More saliently, almost half of the pre-service participants ($n=43$) reported high or very high need for further training to be able to evaluate performance tests (speaking / writing) with a high percentage level 44.4 % which is also a challenging and continuing problem in foreign language teaching and learning as well as assessment process.

When it comes to the qualitative data results, Interview questions 1 and 2 in “Assessment needs” category sought to elicit respondents’ perceptions with respect to LAL specific needs under knowledge, skills and competences domains. The most stated answer was witnessed as practical and theoretical knowledge with regard to alternative language assessment due to the lack of practice.

Research Question 2: What are the language assessment literacy needs and levels of in-service language teachers?

The second research question attempted to identify LAL levels of in-service EFL teachers with regard to language assessment topics to which 39.8 % ($n=39$) of in-service EFL teachers considered themselves to have poor or fair level while 51.0% ($n=50$) of them identified themselves as having good level. Similar to pre-service EFL teachers, in-service teachers also thought themselves “excellent” (38.8 %) with regards to “deciding what to test”. However, associated with “using statistics”, 18.4% ($n=18$) informants stated that they

feel unprepared and unknowledgeable in terms of benefiting statistics to report assessment results. Moreover, “history of language testing” is another topic the participant EFL teachers find their level poor or fair with 28.6 % ($n=28$).

Furthermore, in connection with LAL needs, the second research question revealed that 68.4% ($n=67$) of in-service EFL teachers rated themselves as having a little training, thus being in “very high need” category in terms of LAL related topics. Notably, “Moderate Need” represented the modal response with 48.0 % ($n=47$) of informants. In total, 31.7 % ($n=31$) of in-service EFL teachers either perceived their need “high need” (27.6 %) or “very high need” (4.1%). Thus, only 20.4% ($n=20$) of participants were left in the ‘low or no need’ category. When analyzed in detail, “interpreting scores” and “rating performance tests” seem upfront which were addressed in a very high need category for further training. This might be explained by the fact that the lack of experience of in-service teachers in assessing speaking and writing as a test taker or as an assessor. Because the examination system in Türkiye as an EFL setting does not necessitate speaking or writing competency in the selection of future teachers in university entrance exam (YDS).

When it comes to the qualitative data results, Interview questions 1 and 2 in “Assessment needs” category sought to elicit respondents’ perceptions with respect to LAL specific needs under knowledge, skills and competences domains. A common view was shared by many interviewees with respect to necessary knowledge for online language assessment which was experienced by them in an emergency online education as a result of pandemic (COVID-19) which resulted in a feeling of inefficacy of teachers. That’s why they stated a need for updating themselves to keep up with the recent changes in assessment.

Research Question 3: What are the pre-service and in-service language teachers’ beliefs about language assessment?

Research Question 3 aimed to determine respondent views associated with language

assessment beliefs in general sense through 12 interview questions which investigate the reasons for assessment, good and poor assessment practices, and students' responsibility in assessment among others. To reach an overall understanding of assessment beliefs of participants, each question will be summarized with its main findings. To start, the main reasons for assessment were stated as improvement, diagnosis, making decision, motivation and reflecting. Thus, the main purposes of assessment imply a learner-oriented assessment perspective of participants. In terms of student responsibility, almost all participants addressed an active role for the students by discussing their participation in reflecting on their performance, employing self and peer assessment, collaborating and interacting when necessary. In terms of the impact of assessment on both students learning and teaching practices, the overall picture portrayed notifies that providing feedback, creating an awareness of their weaknesses and strengths, increasing motivation were the main themes as positive effects while stress stood the model response for the negative impact of assessment. Concerning teaching practices, impact of assessment was observed in the comments as diagnosing (identifying lacks), reflecting and arranging future lessons. A different point of view reported by a pre-service participant was a decrease in job satisfaction as another influence of assessment. That's why the participants self reported to use assessment results for learning, instructional and reflective purposes in order to enhance learning by giving feedback, identifying lacks, organizing and arranging the following lessons.

When it comes to the basic features of good and poor language assessment practices, the most salient results might be summarized as valid, reliable, process-oriented and authentic are the main adjectives describing a good and meaningful language assessment while lack of reliability, validity and purpose of a test and lack of language assessment knowledge of teachers are the most noted items for a poor assessment practice. Having identified a good language test, the participants were invited to share their opinions on the most useful assessment methods-tasks which were reported as formative assessment, alternative assessment and portfolio assessment which suggest assessment for learning approach. However, more saliently as a reason for not employing these types of assessment widely, in-service EFL interviewees self stated the following concerns such as

inadequate time, class size, limited knowledge for alternative assessment tasks, heavy workload, to name a few.

Research Question 4: What is the contribution of language testing and assessment (LTA) course offered in pre-service language teacher education programs?

Research Question 4 inquired about respondent perceptions associated with the contribution of LTA course to their LAL development. As reported in chapter 4, the received trainings of two groups of participants differ in that 49.5 % of pre-service EFL teachers reported having extensive training while the majority of in-service EFL teachers (68.4 %) noted a little training for the same question. However, 51.0% of in-service EFL teachers still think that their LAL level is “good” similar to 46.4 % of pre-service EFL teachers. Thus, this might be explained by the fact that classroom experience facilitates and enhances language assessment knowledge and practices of teachers even though a lack of training exists. On the other hand, when further training needs for LAL development are concerned, 68 of the pre-service EFL teachers (70.1%) and 74 of in-service EFL teachers (75.6%) considered their needs either moderate or high.

Research Question 5 and 6: What are the language assessment-related topics that participants recognise as important for their language assessment literacy development? & What are the most challenging concepts of language assessment literacy topics?

Research Question 5 was posed to explore the perceptions of participants with regard to the importance level of LAL topics while the 6th research question investigates the challenging topics for LAL development. As the results indicated, concerning respondents’ perceptions connected to the importance level of LAL topics, 80.4% of pre-service teacher respondents identified “validity” as essential. Remarkably, “reliability”, “interpreting scores”, and “deciding what to test” are the other topics which were represented by the modal response “essential” 72.4 %, 70.4 % and 67.3% respectively in in-service EFL teachers group. On the other hand, pre-service EFL interviewees self-

reported the following LAL topics as both useful and challenging “validation”, “reliability” and “writing test specifications/blueprints”. One other eye catching result was that in-service EFL teachers self stated that they give the least importance to “history of language testing” ($M=2.98$, $SD=1.268$) which agrees with the pre-service EFL teachers’ results with the same item ($M=2.80$, $SD=1.213$). However, this finding contradicts with the qualitative evidence collected through IQ11 which revealed that history of language testing was regarded as “useful” and “important” as it was believed to be a useful way of learning the evolving stages of a language assessment method and/or technique along with its pros and cons, thus enabling teachers to choose the best one fitting their purpose.

The general tendency of interviewees’ thoughts about the challenging topics in LAL development showed that “Validation”, “Interpreting the scores”, “Rating performance tests (speaking/writing)”, “Test analysis”, and “Use of statistics” were the main topics both groups of participants find difficult to cover or apply in an effective way.

Research Question 7: What are the possible ways for language assessment literacy development?

This research question attempts to explore the main ways for LAL development which differ among interviewees. Therefore, the in-service EFL teachers’ responses and pre-service EFL teachers’ opinions seem in alignment which can be summarized and exemplified as follows: academic activities (such as seminar, conference, and workshop), reading relevant literature, professional development programs, and further graduate education (such as Master and PhD programs). On the contrary, a pre-service EFL interviewee suggested self reflection and experience through observation of colleagues and other similar contexts might also foster LAL development of stakeholders.

Research Question 8: What are the pre-service and in-service language teachers’ perceptions about online language assessment?

The last research question is related to online language assessment (OLA) which is inevitable in today's modern world particularly after the global pandemic, thus the responses and views of the participants are crucial for managing the online language assessment process with its expected and unexpected consequences. Initially, OLA was considered as inevitable with its all drawbacks and benefits. A long list of challenges and problems were noted by the interviewees as academic misconduct, difficulty of proctoring, open to manipulation, validity and reliability concerns while practicality, having anywhere anytime opportunity, fun and stress-free nature were the positive sides of OLA based on the responses of participants.

5.3. Conclusions

As it is highlighted by many leading figures in the literature, language teachers' limited understanding of assessment skills (Gardner & Rea-Dickins, 2001; Volante & Fazio, 2007) in addition to poor graded language tests (Alderson, 2005) reveal that there is an urgent need to investigate the pre-service training for language assessment literacy development in all possible angles. Therefore, as the main purpose of the present study was to uncover language assessment literacy needs, levels and beliefs of EFL teachers and teacher candidates with an attempt to question the contribution of pre-service language teacher training programs -language testing and assessment related course in particular- to the LAL development of prospective teachers, the main outstanding conclusions drawn will be put forward in accordance with the abovementioned purposes.

That's, the main framework for presenting the conclusions will be outlined associated with "pre-service training", "LAL development", and "language assessment beliefs" by discussing each research question under related argument.

5.3.1. Conclusions Associated with Pre-service Training

The main conclusions with regard to pre-service language teacher training were drawn from the results obtained through RQ4, RQ5 and RQ6 in addition to the IQ3 in

“Needs of Assessment” section and IQ11, IQ12 in “Assessment Beliefs” section. In this sense, the results and discussion of RQ4 and IQ3 yielded insightful conclusions with respect to the contribution of language assessment related course to LAL development of prospective teachers in general sense. What’s more, the data gathered by RQ5, RQ6 and interview questions -IQ11, IQ12- enabled reaching conclusions associated with the content of LTA course served at pre-service language teacher training programs.

Initially, when data related to the contribution of language assessment related course to LAL development is concerned; the results display a split tendency with in-service EFL teachers arguing against and pre-service EFL teachers arguing for the contribution of the course. This was also supported by the results of RQ1 and RQ2 which uncovered the LAL levels and needs of participants. As discussed before, in-service EFL teachers claimed that the LTA course had not met their LAL needs as it was just an introductory and theoretical, thus considering their training in “a little training” category which corroborates the following studies discussed in the literature such as Chung and Nam’s study (2018), Hasselgreen, Carlsen and Helness’s study (2004), Vogt and Tzagari’s study (2014), to name a few. On the other hand, pre-service EFL teachers in the present study regarded their training as “extensive” because they thought that they have had a great deal of theoretical knowledge. On the contrary to what pre-service EFL teachers have stated, in-service EFL interviewees regarded their training as “limited”.

It can be noted that even though in-service EFL interviewees generally found themselves competent (10 out of 14) in terms of language testing and assessment thanks to their educational background, they reported some deficiencies for language assessment literacy related items and further they have also criticized LTA course stating its being too theoretical and limited. This situation draws attention for the researcher as the dilemma here (considered themselves as competent but feeling incompetence in LAL items) raises some concerns. Firstly, it might be derived from the “social desirability response bias” effect (Tao, 2014: 230). Secondly, and most importantly, practical knowledge attained via their teaching experience and collegiality is possibly the main factor which ensures a feeling of efficacy in terms of language assessment while theoretical knowledge they self

reported to lack of is the main reason for their feeling of incompetence. In this sense, the question that is to be posed here should be 'how to bridge the gap between the practical knowledge and theoretical knowledge' which is not a new story.

Whilst it might be the case that pre-service EFL teacher candidates are unable to predict the possible challenges, problems in real life context as they haven't had opportunity to apply what they have covered in theory into practice. Similarly, providing a firsthand experience, classroom practice enabled in-service EFL teachers to reach a clear understanding of the difficulty of putting theoretical knowledge into practice. The gap between theory and practice is an old story which is voiced by many researchers. But, it is worthwhile to seek solutions to minimize the effect. That's why the following conclusions draw attention to the issue.

First conclusion drawn from this discussion is the unawareness of pre-service EFL teachers about the LAL practices and possible challenges in real classroom settings assuming their training as extensive enough in theory. Öz and Atay, on the other hand, acknowledged that "assessment literacy depends upon teachers' being able to use their assessment-related knowledge effectively in their contexts" (2017: 27). Therefore, their misleading tendency raises concerns for the researcher as assuming they have had enough knowledge may result in not seeing their needs, weaknesses or strengths in terms of LAL knowledge and practices. As a result, awareness-raising practices are suggested which ensures reflective environment in which teacher candidates will have the opportunity to carry out language assessment tests and tasks to various target groups in different assessment methods and techniques whilst simultaneously providing experience-based knowledge for the prospective EFL teachers. In the same vein, as DeLuca and Klinger emphasized an approach is needed which combines practice, theory and philosophy together in assessment education (2010). This could be fulfilled through their practicum courses in which they could meet different groups of learners which create a great chance for practice by producing various language tests and tasks, thus putting the theory into practice. This might be achieved through "apprenticeship approach" to LAL development proposed by Fulcher (2020) in which students are regarded as "apprentices who learn by doing: using theory to design, research to create, values to assess" embedded in learning oriented approach (Fulcher, Panahi & Mohebbi, 2022: 52). Thus, "doing" as a social

activity yields a collaborative environment in which learners as apprentices and teachers as mentors cooperate, interact and reflect (Fulcher, 2020). Therefore, in order to bridge that gap, pre-service teacher trainers should attempt to provide opportunities for future teachers in order to improve their experiential knowledge in collaborative and interactive way which further enhances their language assessment skills either as a test developer and / or as an assessor. This approach towards pre-service teacher training should be also target oriented which to some extent requires context specific knowledge. For instance, assessing young learners' language proficiency in an online language setting is significantly different from assessing EAP learners' language competency as a result of the certain differences of test purpose, test method, test constructs, test takers, score meanings, among others. To this end, as discussed before in results part, language assessment and testing course in pre-service teacher training should cover test purpose (why to test) and writing test specifications (what to test and how to test) in detail by leaving enough room for practice for prospective teachers to handle the abovementioned differences in various contexts in each of which certain LAL competencies are necessitated to produce sound and meaningful language tests and tasks.

One of the other possible explanations for the discussion above with regards to the gap between theory and practice might be expressed through Bloom's taxonomy in which learning is categorized under six cognitively described levels to name "remember, understand, apply, analyze, evaluate and create" (Anderson, Krathwohl, & Bloom, 2001). To exemplify, the interviewees in the present study self reported to have enough knowledge about 'deciding what to test' but limited knowledge and / or difficulty about 'validity, reliability, interpreting scores' to name a few. In this sense, according to the Bloom's taxonomy, the first two stages – remember and understand- are not problematic for the participants' LAL knowledge and competency; however, the upper levels in the taxonomy which requires implementing, drawing connections, justifying and producing are the ones raising serious concerns. With regards to language assessment, these levels bear greatest importance as they are the very steps to create a meaningful assessment task. To start, writing test specifications consisting of test purpose, sample test items, scoring for each item, etc. is one of the main ingredients for LAL development which necessitates making connections between test purpose and test method, test construct and test item, test

item and test score, test score and feedback, among others. That's why in order to draw such connections, a pre-service teacher candidate should have enough theoretical knowledge and practical knowledge as well. Secondly, in association to evaluating stage of Bloom's taxonomy, selecting a test to use for your own context or developing a new one requires justifying the test method and the underlying paradigm with a specific focus on test purpose as well. Lastly, creating a new test requires to deal with washback, ethical considerations, social context, administration among others to formulate a new language assessment test or task through a critical eye, awareness and reflective point of views. Therefore, while designing and presenting language assessment course at pre-service teacher training, it would be beneficial to take into consideration these cognitive levels so that teacher candidates can build a bulk of knowledge in both practical and theoretical aspects.

One other conclusion drawn from these findings can be the fact that in-service EFL teachers developed their main language assessment skills in their profession which is also attributable to the positive impact of collaborating with colleagues. Consequently, as a researcher this raises concerns for me that collegiality serves as a means of enhancing LAL skills and knowledge. Therefore, it would be useful to create an opportunity for teacher candidates to observe, evaluate, and analyze the language tests used in real classrooms in order to build necessary related knowledge. This can be also explained by Vygotsky's theory of the zone of proximal development in which pre-service teacher candidates can do their own and they can fulfil their potential with the help of in-service teachers which further assists in fostering LAL.

Moving the discussion further, based on these salient conclusions drawn from the findings above it would be beneficial to implement a more promising teaching strategy "inappropriateness analysis" to equip future teachers with necessary knowledge about test development, test specifications and evaluating a test. The analysis of inappropriate tests and/or test items might be an important way to foster language assessment literacy by encouraging critical thinking and learning. As a coined word, "inappropriateness analysis" is purposefully termed by the researcher of this present study as "inappropriateness" suggests that a test may be designed well without any mistakes or errors but may not be fit for purpose which causes the test to be "inappropriate". That's why mistake analysis or

error analysis is not found suitable to explain this way of learning. In “inappropriateness analysis” technique, one can reach an understanding of what is not acceptable in a language test by taking test purpose, target population, test taker, test context, ethical concerns, to name a few into consideration. By doing so, one can distinguish between a good and a poor test which requires analysing the foreseen or unforeseen effects of the test, possible problems for test takers, learning process, and teaching instruction, etc. Additionally, one can have a chance to explore the possible reasons of badly designed tests by determining the source of the problems (e.g. face validity, inappropriate test method, unreliable test scores, among others) which further assists in developing coping strategies to solve the problems. This way of learning including problem-solving and critical thinking skills, in the long term, ensures increased motivation and learning at the same time as pre-service teachers are expected to be able to design a better test by eliminating the inappropriateness, problems and mistakes in a badly designed test which is similar to negative evidence as well. Because of the fact that negative evidence provides information about what is not correct and appropriate in language, inappropriateness analysis of the sample language tests also provides such evidence for the inappropriateness in the process of designing a test, administering a test and interpreting scores. It is not unfair to say that there are a lot of samples for the inappropriate language tests which were widely discussed in the study of Köksal by listing the main concerns aroused within the tests (2004). Through inappropriateness analysis, minimizing the probability and frequency of occurrence of the same poor items or tests might be achieved thanks to learning from one’s mistakes which leads to developing a more sound and meaningful language assessment tasks and items.

When it comes to the language testing and assessment course, the data reached by RQ5, RQ6 and interview questions -IQ11, IQ12- provided valuable insights related to the content of LTA course served at most of the pre-service language teacher training programs. The questions in the present study initially attempted to explore the perceived importance levels of LAL related topics given by participants and further to find out the most challenging and useful LAL related topics. The implicit purpose here was to determine the essential language assessment related topics to cover in LTA course at pre-service language teacher education which is believed to assist in building necessary background knowledge for prospective teachers.

A great deal of data including a cluster of responses to IQ11 and IQ12, and quantitative results of Part 3 in the questionnaire showed that “validity and reliability” were the most required topic to cover in LTA course while “history of language testing” was perceived as the least essential subject. Moreover, in-service EFL teachers regarded “deciding what to test”, “writing test tasks and items” and “interpreting scores” as the most essential topics. This might be explained by the fact that in-service EFL teachers are aware of the need to develop skills and competences for test design and development along with interpreting scores which are worthwhile to give constructive and meaningful feedback to students. In this sense, certain competences and qualities are undoubtedly necessary to be acquired by teachers such as selecting assessment methods, developing assessment methods, developing valid student grading procedures, communicating assessment results, and recognizing unethical, illegal and inappropriate methods of assessment which are also reported in the *Standards* (AFT, NCME, & NEA, 1990).

When data linked to the challenging and useful LAL topics is considered, the overall picture seems similar for both in-service and pre-service EFL teachers. The general tendency of interviewees’ thoughts about the issue showed that “Use of statistics”, “Validation”, “Interpreting the scores”, and “Rating performance tests (speaking/writing)” were found as the most challenging topics. This situation raises some concerns for the researcher as these abovementioned topics necessitate developing valid tests, analyzing the scores and communicating the results in an effective and objective way. Therefore, one conclusion drawn from these findings can be the need for further training which would help teachers enhance their knowledge, skills and competences with regards to these abovementioned LAL topics, thus building a satisfactory level of LAL. This might then serve as a means for conducting meaningful language assessment practices, thus supporting learning and teaching in the long term.

As the quantitative data showed “rating productive skills (speaking / writing)” can be challenging for EFL teachers which corroborates qualitative data as well. Whilst it might be case that in foreign language context it becomes difficult to clearly define the construct and test purpose for speaking and writing by answering the question of what counts as evidence for speaking and writing competence. As a result, in pre-service

language teacher training the necessary skills and knowledge with regards to determining test construct and purpose should be dealt with by providing opportunities for practice through guiding model implementations which assist prospective teachers in reaching a clear understanding of the issue with regards to language assessment and testing.

Another possible explanation for this situation maybe the case that EFL teachers are not able to administer and rate productive skills effectively as they are themselves in lack of experience as a test taker, because the examination system for selection of future EFL teachers in Türkiye is grammar based including multiple choice questions in the exam of YKS and YDS. In this sense, one other conclusion which can arguably be drawn from this discussion is the challenges and difficulty in ensuring validity and reliability in administering and rating productive skills can be attributable to the lack of experience of EFL teachers in Türkiye as a test taker. This corroborates the results to RQ7 and IQ11 and IQ12 as well. Because of the fact that these abovementioned results showed that the majority of the participants perceived “reliability, validity and interpreting scores” as the most important topics which are the most challenging concerns in rating and administering productive tests which all require an objective perspective, discourse knowledge and competence, communicative competence particularly in the case of speaking, among others. Needless to say, these competences are context-driven and situation-driven in nature. Thus, the multifaceted and contextualized LAL come into play here which is worthwhile to consider while designing the language assessment course in pre-service teacher training in order to assist prospective teachers in conducting their assessment practices effectively.

In doing so, the main tendency should be to equip future teachers with necessary practical knowledge in terms of language assessment related topics so that language assessment practices in real classroom settings will be meaningful and applicable. To this end, prospective teachers should be given enough opportunity to put their theoretical knowledge into practice in language assessment course during their pre-service education. For this purpose, the prospective EFL teachers may be grouped as test developers or test takers during their language assessment course in which experiential knowledge should be enhanced. To illustrate, in this pretend-to-be teaching strategy, prospective EFL teachers

grouped as test developers may be required to produce sample language tests some of which might have content validity but lacks face validity or they might produce some sample tests for different target groups such as young learners, EMI learners, EAP learners, to name a few. By doing so, they may witness potential challenges and problems in each context and further become aware of their weaknesses and strengths in order to develop certain coping strategies in real life. On the other hand, the primary benefit of pretending to be a test taker is that prospective EFL teachers as test takers are able to understand the feelings and expectations of learners along with witnessing the effect of the test (washback) on their learning at first hand.

For these purposes, LTA course at pre-service language teacher education needs to be paid attention in order to design an effective language assessment course including the objectives, outcomes, content, suitable materials along with the timing (8th term) and weekly course hours of the course with an attempt to develop LAL of prospective teachers. Remarkably, Volante and Fazio underlined in their study that a specific course in assessment is necessary to improve assessment knowledge (2007). However, the main difficulty lies in determining a suitable content (what to teach), the methodology to use (how to teach), and the underlying philosophy (why to teach). Therefore, one of the major contributions of this present study is believed to be a fruitful discussion about LAL topics which would allow different parties (program developers, policy makers, teacher trainers, teachers, teacher candidates, etc.) to better conceptualize a language assessment course by throwing shed light on “what to cover, why to require and how to acquire” questions for pre-service teacher education programs which have been widely discussed in this present study.

5.3.2. Conclusions Associated with Language Assessment Literacy Development

A cluster of responses gathered through RQ1, RQ2, RQ7 and IQ1, IQ2, IQ4, IQ5 in “Needs of Assessment” section ensured valuable conclusions with respect to LAL development of EFL teachers and teacher candidates.

Initially, a comparison of the received training and perceived LAL levels were demonstrated based on the quantitative data gathered to discuss RQ1 and RQ2. As it was witnessed, the ones in pre-service group with little training regarded their level as average while extensive training receivers considered their levels in good category. In similar vein, one half of the in-service informants who reported having a little training regarded their level as good or fair. This arouses attention that having a high self efficacy level in terms of LAL may be attributable to having taken the course just at the time of the data collection, that's recently acquired knowledge about language assessment including basic concepts and terms have probably ensured higher self confidence. This finding corroborates Öz and Atay's study (2017). Another possible explanation for pre-service EFL teachers' overrating their LAL levels might be attributed to "social desirability response bias" as well (Tao, 2014: 230). When in-service data considered, almost one third of teacher respondents regarded their LAL related knowledge and skills in fair level which might be explained by the fact that classroom experience enabled to monitor their assessment practices, thus being aware of the main weaknesses and problems. As a result, practicum course at the pre-service teacher training is noteworthy for making the best use of it by ensuring practice opportunities for future teachers which is believed to enhance LAL and language assessment practices.

The data linked to the needs for further training in relation to LAL was reached through Part 2 in the questionnaire and "Needs of Assessment" part in interview session. Consequently, almost half of the participants in each group actively acknowledged their need for further training as "moderate" no matter what their level is and no matter how much training they have received. This raises concern for the researcher that theoretical knowledge is helpful to some extent, yet still it is crucial to equip teachers with experience based knowledge in order to effectively manage multifaceted and complex nature of LAL development.

Having discussed the conclusions with regard to the training received and LAL needs, it is noteworthy to draw conclusions about the main ways for LAL development participants self reported. The findings of RQ7 apparently reveal that academic activities

(such as seminar, conference, workshop, post graduate studies, etc.) and professional development programs (such as in-service trainings) are predominantly the most adopted ways for LAL development. This was also supported by results of IQ4 and IQ5 which displayed qualitative data including emergent themes which were found as personal experiences, classroom practices, peer observation, to name a few.

Based on these clusters of comments and quantitative data, a framework for designing a LAL development program has been developed by the researcher herself which is believed to be one of the major contributions of this present study to the field. When supported with empirical studies including research-driven-data to validate the framework, it will be a guiding road map to design a LAL development training program by considering philosophical, practical and developmental aspects.

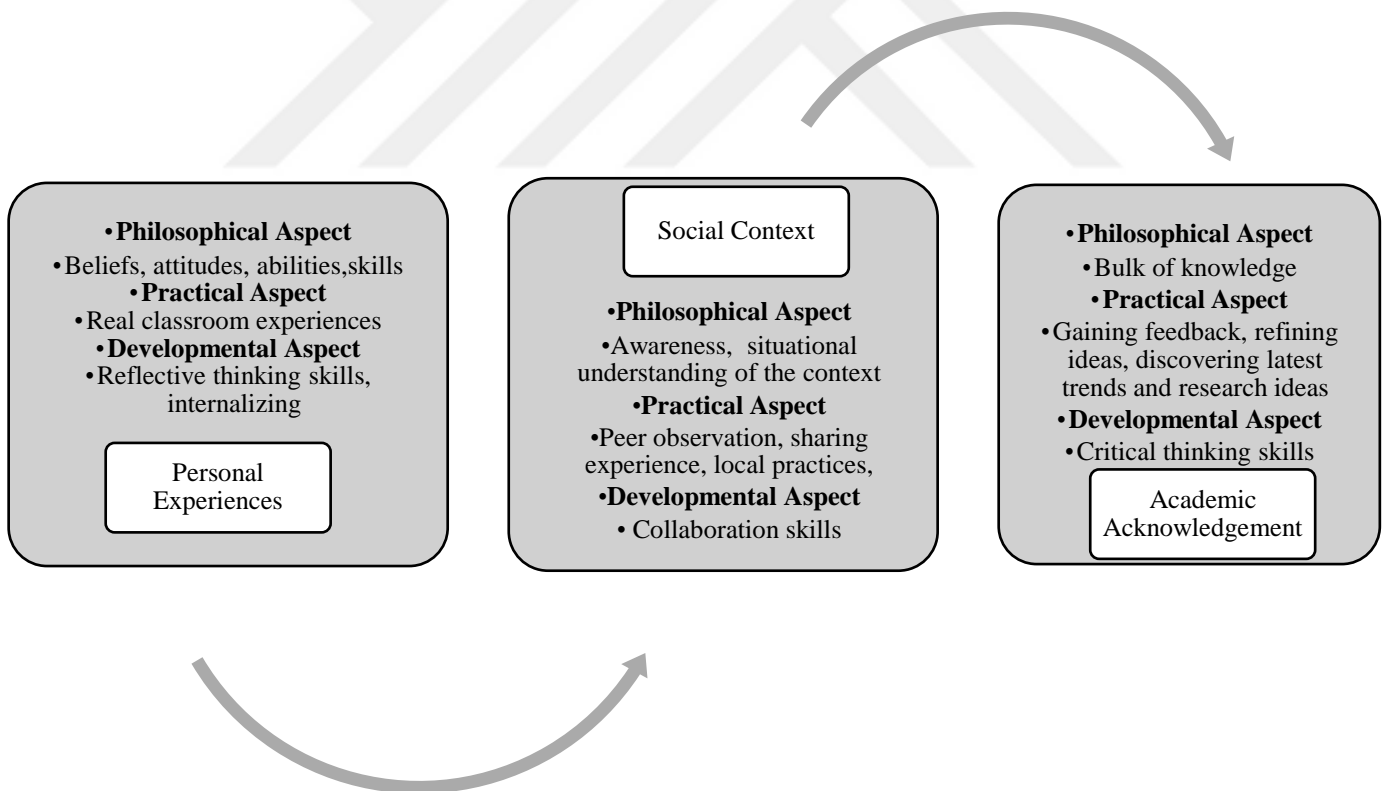


Figure 5. A framework for language assessment literacy development

As it is displayed above in Figure 5, the LAL development framework is based on three main domains named as “Personal Experience”, “Social Context” and “Academic Acknowledgement” each of which consists of three layers as follows: Philosophical Aspect, Practical Aspect and Developmental Aspect. The mentioned framework has been already detailed and discussed in Chapter 4. Therefore in this part, only the related interpretations and conclusions drawn will be shared.

To start, it is quite possible to conclude that academic knowledge built through post graduate studies, conference/seminar works mainly by individual effort is one of the key considerations for LAL development. A concern is raised here for the researcher that the best ways for LAL development requires gaining feedback, refining ideas, discovering best practices, collaborating, reflecting among others. All these activities are related to critical thinking and reflective thinking skills which should be fostered in order that LAL development can be ensured and the quality of language assessment practices can be enhanced.

It is advisable to consider widening “Collegiality” among language teachers and teacher candidates as reflective thinking and collaborative activities foster language assessment literacy by creating opportunities to explore best practices through peer observation and feedback. This also corroborates “contextualized, collaborative, and reflective” nature of LAL as put forward by the following figures as contextual, experiential concerns (Yan & Fan 2020), self-reflection and apprenticeship (Scarino, 2013), collaborative and reflective assessment practices (Harding & Kremmel, 2016; Levi & Inbar-Lourie, 2020; Vogt & Tsagari, 2014) for LAL development. In line with this, Yan and Fan (2020) argued for apprenticeship-based, experience-mediated model. To design such a model, the framework developed in the current study would allow program developers to consider personal experiences, social context and academic acknowledgement taking philosophical aspect, practical aspect and developmental aspect in each step into account. For example, while designing a LAL training program, with regard to social context, it is quite necessary to comprehend the importance of the context which embodies the setting, the people, the language policies, etc. embedded there.

Therefore, for the first stage-philosophical aspect-, being aware of the abovementioned issues (setting, people, policy, etc.) and situational understanding of the context are the main ingredients for a successful LAL development. For the second stage, some certain activities such as peer observation or sharing experience could be embraced in order to improve language assessment practices. Moreover, the last stage requires collaboration skills to make use of peer interactions and experiences. To illustrate, as proposed by Yan and Fan (2020), self-reflection and apprenticeship also might facilitate LAL development in addition to contextual and experiential tenets. More comprehensively, “mentors, colleagues and published assessment materials” (Tzagari & Vogt, 2017: 52), “exchanging knowledge and experience, observing others’ assessment practices, asking for colleagues’ guidance” (Babaii & Asadnia, 2019: 757), reflections and assessment practices (Levi&Inbar-Lourie, 2020) can be regarded beneficial ways for LAL development which were also discussed in this current study.

5.3.3. Conclusions Associated with Language Assessment Beliefs

Results mainly from RQ3 (data obtained through ten interview questions) and RQ8 provided a wealth of information about language assessment beliefs of participants which further shed light on the following conclusions.

In general sense, the main findings reached through ten interview questions will be grouped under five related arguments in order to draw and present salient conclusions. These arguments are conceptualized as language assessment purposes, qualities of good-poor language assessment practices, students’ responsibility in language assessment process, the effects of language assessment practices, and finally perceptions of online language assessment.

Initially, a great deal of data obtained by IQ1 and IQ4 revealed that the main language assessment purposes participants reported to have as “making decision”, “improvement”, “diagnosis”, “motivation”, “reflection”. This finding was also supported by results from IQ8 which showed the majority of informants support assessment for

learning. Needless to say, it can be observed that the main purposes of assessment were considered as enhancing the quality of education and teaching instruction, identifying the problems, strengths and weaknesses within the teaching and learning process. It is possible to conclude that the general tendency is towards assessment for learning which emphasizes the process and learner based assessment in which the primary attempt is to foster learning through benefiting from the assessment itself. Another conclusion which can possibly deduced from this perspective is the self-awareness of both in-service and pre-service EFL teachers in this study about the positive sides and advantages of assessment for learning approach by employing alternative assessment and formative assessment methods like portfolio tasks. Therefore, it would be better to equip prospective teachers with necessary skills to carry out such assessment practices in an effective and meaningful way, so that learners can benefit from the assessment itself rather than getting a score.

Having identified the assessment purposes of respondents, now it is highly time to explore in which way they carry out a language assessment to serve abovementioned purposes, so that good or poor assessment practices occur. A pile of comments with regard to IQ6 and IQ7 addressed these concerns. In terms of a good language assessment, the findings revealed the main qualities as “validity, reliability, authentic, process-oriented, and feedback providing”. It is possible to conclude that the participants have a tendency to embrace a learner oriented purpose to monitor the progress of learners in a meaningful way by providing feedback. One concern the researcher calls attention to is the time allotted to such assessment practices which is quite limited due to inadequate time, class size, limited knowledge for alternative assessment tasks, heavy workload, etc. As a result, it is noteworthy considering a change in weekly class hours of teachers which would allow them to plan and administer such alternative tasks in addition to providing in-service trainings for making best use of alternative assessment tests and tasks. When the results of IQ7 are concerned, it is revealed that test-related and teacher-related problems are the main reasons for poor assessment practices. These might occur in the format and content of a test (exp: invalid tests), inappropriate way and/or tasks (stemmed from lack of LAL knowledge of teachers), among others. One inference which can be derived from this finding is the need for further training to enhance the quality of language tests which in turn fosters learning.

The data obtained by IQ9 in order to explore opinions of participants with respect to students' responsibility in language assessment process uncovered the general tendency of participants on attributing an active and reflective role for learners. As a result, one of the main conclusions derived from the comments is the importance of learners in assessment process as an active player who needs to monitor, self regulate and control his/her language assessment. Therefore, just like the learning process, assessment process needs to be managed by the learner herself/himself from the first stage as purpose of assessment to the last stage as interpreting scores. Remarkably, the main concerns the researcher wishes to voice are the willingness, awareness and autonomy which are the key considerations for a learner to self regulate his / her language assessment through which s/he benefits the interpretation of the assessment results. To fulfil these purposes, the learner should have reflective and critical thinking skills to willingly monitor his/her language assessment practice by diagnosing missing parts, weaknesses, strengths and further use that information to enhance his/her learning. Unfortunately, examination oriented system just like in Türkiye, does not allow learners to be a participant in assessment process, thus the learners could not develop their reflective, critical thinking and autonomous skills.

Based on the brief discussion for the first conclusion above, the following conclusion might be regarded as a suggestion as well. Therefore, one other conclusion which can be inferred is the need to integrate language assessment practices and LAL training into other courses at pre-service language teacher education by embracing self assessment, peer assessment, class observations, reflective journals in these courses which in turn helps prospective teachers internalize assessment process as a whole by putting them in the centre both as a test taker and as an assessor. As a result, the prospective teachers take an active role in assessment practices which paves the way to using the same assessment techniques when they become a teacher.

Even though a great amount of time is needed to train students to assess their own and peers' tasks, the responsibility learners take and the process of making decision about the quality of the tasks are the main benefits of self and peer assessment (Fulcher, 2010).

Moreover, certain implicit benefits can be also noted as engagement of learners, responsibility, critical thinking and reflective thinking in terms of assessment of their own works or others', thus enhancing learning and motivation. This corroborates Şahin's study (2018) where she mentioned the positive effect of being engaged in assessment practices for instructors for conducting effective assessment. In the same vein, Fulcher (2010: 69) also emphasised that "the process of assessment, including self-assessment, could improve motivation and self-esteem, leading to additional learning gains". As either the assessment itself or the feedback on the basis of assessment score provides learners with a sense of achievement which in turn encourages them to study more. Therefore, learners need to be aware of the positive effects of assessment and should be encouraged to take an active role in assessment process. However, moving the discussion further, it is quite significant to note here that one of the most salient issues is also 'teacher belief' which is considered to have an important effect on the way teachers regard students' role in assessment process (Barnes, Fives, & Dacey, 2015).

Interview questions IQ2, IQ3 and IQ5 provided valuable insights related to the impact of language assessment on students' learning and teaching instruction along with the use of assessment results. The questions initially attempted to explore the main effects of language assessment on learning and teaching and further to find out the ways participants follow to make use of assessment results. In this sense, the results and discussion of IQ2 yielded insightful conclusions with respect to the effect of language assessment on student learning described as "motivation", "awareness" and "feedback". In this regard, it is quite possible to conclude that the positive washback effect of language assessment is noteworthy to consider as it ensures a facilitative impact on learning by creating awareness, providing feedback, diagnosing weaknesses and strengths, increasing motivation and participation, fostering learning, among others. More saliently, in association with feedback, the main perspective of the interviewees showed that it is only beneficial for learners if it enhances their learning. In similar vein, the significant figures from the related study field supported that feedback should increase awareness of learners about what they know, what they need to know to improve their learning and what the next step they attempt to reach, thus understanding their learning goal which in turn fosters self regulated and autonomous learning (Fulcher, 2010; Rea-Dickins, 2006). In this sense,

providing feedback is undoubtedly multifunctional in that it both serves as a meaningful guidance for learners by making them aware of their current level and the target level and as a mediator between learning and learner by providing them opportunity to manage and monitor their own learning. However, stress was another mentioned effect of language assessment on students' learning which deserves a great amount of attention. This negative impact should be taken into consideration while designing a language assessment task and giving feedback to learners, so that the assessment fosters learning through constructive feedback in a stress-free environment.

The data coming from IQ3 which questioned the impact of language assessment on teaching instruction unveiled two groups of impact categorized as “As teaching impacts” and “After teaching impacts”. “Diagnosing (Identifying Lacks)” was regarded as an “as teaching impact” revealing the importance attached to explorative aspect of language assessment impacts by interviewees. “Critical thinking and Reflecting” and “Organizing and Arranging” were considered other impacts of assessment in “After teaching” category. One conclusion which can be induced from these findings is the significance of language assessment in shaping the teaching instruction by taking an action according to the assessment process either as teaching or after teaching. Considering a cluster of remarks on the use of assessment results, IQ5 shed light on the following purposes “Learning”, “Instructional” and “Reflecting”. One main conclusion derived from this finding is the general tendency of interviewees to use assessment results for providing feedback so that student learning is fostered. More importantly, it is crucial to note here that “Scores only have meaning when they’re used for defined purposes that informed the design of the specifications” (McCallum & Coombe, 2022: 15); otherwise, they do not ensure a constructive and meaningful feedback for enhancing student learning and teaching instruction. Additionally, using assessment results for instructional and reflective purposes also suggest that EFL teachers and teacher candidates may benefit from the assessment results to organize and plan their future lessons and further self criticize their own way of teaching. Consequently, these mentioned purposes –also can be regarded as competences– are related to the key considerations in LAL for which there is a need for further training in order to create awareness and build a clear understanding for a sound language assessment.

Moving the discussion further, the impact of language tests upon teaching and learning process has also something to do with preparing learners to take test which leads teachers and students to study for the test not for the language itself (Fulcher, 2010). This is also a true picture for the situation in Turkish examination system which forces teachers to prepare students for the high stakes tests such as YKS, YDS by teaching them test-taking strategies instead of teaching the language. More saliently, there is a hidden agenda behind these tendencies which is embodied by economical concerns including mass production of testing and teaching materials, the prestige of schools, increasing of private lessons (Fulcher, 2010). What's more, the main tendency in exam-oriented education systems is that teachers are expected to prepare learners for the test through test taking strategies or test items and tasks similar to the test itself. Additionally, this kind of teacher is regarded as the best and knowledgeable one by parents, even by institutions as they are believed to increase students' scores in high stakes test, thus raising the school prestige. This raises serious concerns both for the language teaching and also learning process in the long term.

The data linked to the perceptions about online language assessment (OLA, henceforth) was obtained through Part 3 in the interview section consisting of five interview questions to elicit responses on the following issues as overall idea for OLA, suitable tasks and tests for OLA, online feedback, challenges in OLA and coping strategies if any. Consequently, almost all of the participants in each group shared the idea of inevitable feature of OLA. This raises concern that in today's modern world it is noteworthy considering effective ways for designing, administering OLA and interpreting the scores. In pursuit of presenting the opinions explored in IQ4 in "Online Language Assessment" section, a long list of challenges and problems were noted by the interviewees as academic misconduct, difficulty of proctoring, open to manipulation, validity and reliability concerns. One conclusion drawn from these findings is the need for finding effective solutions to carry out OLA in an effective and reliable way. This finding was also supported by results from IQ1 and IQ3 which showed the majority of informants argue against OLA as it is difficult to control and get valid and reliable results. When positive sides of OLA are considered, a good deal of information unveiled that practicality, having anywhere anytime opportunity, being stress-free and fun come to the forefront. In this

sense, it is quite reasonable to contemplate on getting the most out of OLA by organizing in-service training activities in a wider range of contexts and purposes.

Considering suitable online language assessment tests and tasks, the findings bring about three categories as “Higher-order thinking tasks”, “Interactive tasks” and “Traditional tasks”. In each category, certain language assessment tools are voiced by the informants such as individual projects, research based tasks, portfolio activities. One conclusion which can be inferred is the awareness of participants with respect to interactive and higher order thinking activities which would also assist in controlling academic misconduct. In doing so, assessment leads to getting reliable and valid data which in turn provides meaningful feedback for the learner. Overall, the findings related to the challenges in OLA practices revealed that student-related, teacher-related and technology-related challenges which were categorized by the researcher. This situation raises concerns for the researcher as lack of technological knowledge of both students and teachers voiced by interviewees may cause a decrease in the effectiveness and quality of OLA practices. As a result, it is an urgent need to develop in-service trainings with regard to information and communication technology use as ICT literacy has a facilitative impact on online learning and assessment. The main coping strategies were also questioned through IQ4 and reported as proctoring programs (Turnitin, Witwiser, etc.), limited time for each question, distribution of questions, and variety in the types of questions and tasks, to name a few. Therefore it is possible to conclude that these strategies would help teachers minimize the negative consequences of problems and challenges faced in OLA.

5.4. Implications

The implications of the study based on the findings and conclusions drawn will be discussed from three aspects, to name theoretical implications, methodological implications, and practical-pedagogical implications addressing different stakeholders’ interests and needs. In this sense, this current study bears importance as it involves two key stakeholders with regards to their LAL needs and beliefs with a purpose to explore the main language assessment related areas which need further training or improvement. Reaching these informant groups ensures the main strength of the study, because doing so

different contexts are taken into consideration, that's pre-service teacher education setting (EFL teacher candidates) and real life classroom setting (in-service EFL teachers).

As it is widely accepted and supported by many research studies, pre-service teacher education programs are insufficient in equipping the future teachers with necessary practical skills, basic competences and knowledge for LAL. With this insufficiency in mind, policy makers, to name CoHE should identify the objectives and outcomes of language assessment and testing course in a collaborative manner with the ones who are responsible for teaching this course -teacher educators-, with the ones who are the end users of the course -in-service EFL teachers-, and with the ones who are active participant in the teaching and learning process -pre-service EFL teachers-. In doing so, the course content armed with useful practical skills needed for real life classroom and instructional purposes will help future teachers enhance teaching and student learning through assessment.

5.4.1. Practical and Pedagogical Implications

With an attempt to compensate theory-practice gap in terms of language assessment practices, LAL development of student teachers bear greatest importance which needs to be taken into account by teacher educators who are responsible for language assessment course and further teacher education program developers who are the main stakeholders behind the rationale of the course content. Based on the discussion here, the stakeholders responsible for making decision about the content, teaching-learning process, objectives and assessment of language assessment course at pre-service teacher education programs should have at least an acceptable amount of LAL knowledge and skills which in turn facilitates improvement of the course.

Moreover, improving LAL of prospective EFL language teachers is generally ensured through one specific language assessment and testing course. However; as Şahin (2018) shared, as a result of insufficient time, to cover necessary language assessment related topics within one single course was stated as one of the main challenges for

instructors. Further, when inquired closely, it can be seen that communicative language competence is strongly advised by Ministry of National Education (MoNE) which requires assessing language from such domains as communicative language testing, critical language testing, culture sensitive language testing, etc. Nonetheless, it is undoubtedly accepted that pre-service teacher candidates do not feel competent to administer effective assessments, let alone abovementioned aspects. As Öz and Atay acknowledged that "assessment literacy depends upon teachers' being able to use their assessment-related knowledge effectively in their contexts" (2017: 27), awareness-raising practices might be helpful which ensures reflective environment in which teacher candidates will have the opportunity to review existing language assessment tests and tasks or develop their own tests for various target groups in different assessment methods and techniques whilst simultaneously providing experience-based knowledge for the prospective EFL teachers. This could be fulfilled through their practicum courses in which they could meet different groups of learners which create a great chance for practice by producing various language tests and tasks, thus putting the theory into practice. This might be achieved through "apprenticeship approach" to LAL development proposed by Fulcher (2020). Therefore, in order to bridge that gap, pre-service teacher trainers should attempt to provide opportunities for future teachers in order to improve their experiential knowledge in collaborative and interactive way which further enhances their language assessment skills either as a test developer and / or as an assessor. To this end, language assessment and testing course in pre-service teacher training should leave enough room for practice for prospective teachers to to apply what they have learned during LTA course which enables them to see their weaknesses and needs to produce sound and meaningful language tests and tasks.

For these reasons, based on the main findings and the most salient conclusions of this current study, a practical and pedagogic model (see Appendix 5) based on "apprenticeship approach" Fulcher (2020) was proposed by the researcher for guiding and informing teaching and learning of language assessment literacy of pre-service EFL teachers undertaking LTA course during their pre-service language teacher training. As an explicit pedagogy, this model provides a technique and/or principle to name inappropriateness analysis along with a detailed guideline including content, materials-resources, teaching-learning strategy, activities, roles and actions, learning outcomes and objectives, feedback and assignment which could be used to teach related LAL topics.

In terms of LAL development of in-service language teachers, on the other hand, some practical implications are worth of mentioning. Based on the respondents' views with respect to LAL development, the figure designed by the researcher (Figure 4 and 5) to draw a main framework might be used to describe the key considerations in LAL development process. Initially, because of the fact that collaborating and sharing experiences contribute to LAL development, thus both pre-service and in-service teachers should be encouraged to join professional development programs which could be helpful for raising awareness of LAL, getting necessary adaptation skills to modify assessment tasks suitable for their own contexts, learning from each other, reflecting and collaborating, among others.

5.4.2. Theoretical implications

Although, discussed as one of the conceptual limitation, an identified framework for LAL competences and skills is not present, still then as the purpose of this study was to determine LAL needs, levels and beliefs of the participants; theoretical framework in this current study could be potentially useful in conceptualizing LAL competences in other research studies. To this end, context sensitive and flexible LAL items need to be conceptualized which could be useful and adaptable to different settings and individuals.

5.4.3. Methodological Implications

The main purpose of the study is to find out LAL levels, needs, and beliefs of different stakeholders from their own lenses. In the pursuit of this aim, 8 research questions were posed through which the researcher has investigated the issue both qualitatively and quantitatively. Therefore the findings of the study based on the self responses by two different parties (EFL teacher candidates and EFL teachers) might create new research avenues. To start, one possible future direction of research could be to explore alignment or discrepancy between language assessment beliefs and language assessment practices through observations and reflective journals. Moreover, such exploration offers insights into the underlying reasons for theory and practice gap which need further investigation.

What's more, in today's technology world, the integration of technology into language assessment tasks is of vital importance as it has been witnessed obviously during the Covid19. Thus, future teachers need to be aware of technology mediated LTA, the main applications, the possible challenges in addition to coping strategies to overcome the problems. Another methodological implication is concerned with suitable language assessment methods and tasks for young learners. Thus, the results of the study may assist in designing a LTA course with appropriate topics involving language assessment of young learners.

This study yielded a clarification to the need for renewal of LTA course content which includes clear and attainable objectives, knowledge and competences for developing suitable language assessment tasks in real classroom settings, the basic skills for being able to assess particularly productive skills (speaking and writing), employing technology mediated language assessment, being aware of ethical concerns in an assessment task, and alternative assessment methods among others.

5.5. Limitations of the Study

The limitations of the study will be discussed in terms of conceptual limitations, methodological limitations, and research design limitations.

5.5.1. Conceptual Limitations

As for conceptual limitation to the present study, the multifaceted and complex nature of LAL poses a challenge for identifying a precise definition of the term and relevant competences. Due to the lack of a clear definition and competences, not all possible of LAL related skills, knowledge and competences are included in this study thus having a risk of missing some skills, knowledge and competences with respect to LAL. Even though the study attempted to give place outstanding LAL models and some

definitions (Fulcher, 2012; Taylor, 2013), still then it is possible to state that some LAL related knowledge, skills, and competences might be probably missing.

5.5.2. Methodological Limitations

When it comes to methodological aspect of the study, collecting an adequate and reliable data was rather challenging as a result of the unexpected outbreak of pandemic. Therefore, data collection process was interrupted two times (2020 Spring Term, 2021 Spring Term) and at least two years stringed out from the planned time. Not only the time loss but also potential data loss occurred due to the transition to online education. Having all these in mind, the study included a limited number of participants even though they consist of two different stakeholders, to name pre-service EFL teacher candidates and in-service EFL teachers. Concerning the sampling, even though purposive and convenience sampling was consulted, still the heavy workload of participants and negative impacts of Covid19 made it impossible for some participants to take place in interview sessions, although initially they all had agreed to voluntarily participate.

For quantitative data of the study 194 respondents (96 pre-service EFL teachers, 98 in-service EFL teachers) filled out the questionnaire while the qualitative data was gathered from 14 in-service EFL instructors and 2 pre-service EFL teachers via semi structured and focus group interviews. Another methodological limitation of the study is that the pre-service EFL teachers have been studying at one state university, that's why qualitative data obtained from students may not be generalized to other pre-service EFL student teachers. Although the generalization of the data is not attempted in the first place in a qualitative research design (Dörnyei, 2007), nevertheless reaching a big number of participants would ensure richer insights and a deeper understanding of the phenomenon. That's why portraying a general overview for the issue under examination in the current study is questionable based on the limited number of informants, still then the responses with regards to LAL levels, needs and language assessment beliefs provided by the participants lead to practical implications in order to tailor assessment related courses offered at pre-service language teacher education.

Considering data collection instruments; course documents, open ended interview questions and a questionnaire were utilized, however they were not capable of including all aspects of LAL because as a researcher I had to take some concerns into account such as fatigue threat because of excessive time for data collection process. As a result, the instruments have a limited number of questions investigating certain aspects, undoubtedly missing some parts.

One potential goal of such a study could be to portray a wider perspective for LAL development during pre-service language teacher education. To throw shed light on the complex and multi dimensional nature of LAL, perceptions of different stakeholders bear greatest importance, in doing so research studies with a considerable number of participants would undoubtedly yield more valuable and reliable data.

5.5.3. Research Design Limitations

As a mixed methods research design, the current study has already endeavoured to produce promising results by ensuring triangulation in terms of data source, data analysis, and data type (Creswell, 2009; Lincoln & Guba, 1985). Within this scope, the quantitative data of the study was gathered through LAL questionnaire and included responses from 98 in-service EFL teachers, 96 pre-service EFL teachers, thus a total of 194 informants. The obtained responses were used to explore the LAL levels, needs, and perceived importance levels of LAL related topics. To fulfil the main purposes of the study, some statistical measurements were employed including frequencies, percentages, and descriptive analysis, and further cross tabs where necessary. Concerning the qualitative part of the study, the data was collected through interviews only. As it is widely accepted, qualitative data could be gathered through observations, reflective journals, diaries, field notes, etc. Lack of such data collection tools would cause another limitation to the current study. Moreover, even though a considerable amount of qualitative data was obtained, obviously it is not reasonable to mention all quotes of participants, hence a limited number of interpretations were included within the study.

In sum, researchers should undertake a compromised endeavour to overcome such limitations in future studies to be able to review the issue more critically, by capturing a comprehensive understanding.

5.6. New Research Avenues for Future Studies

The discussion about the results for each research question has raised some concerns which would be reasonable to take into account in future research. Within this scope, one new research avenue could be the language assessment practices in real classroom settings with LAL in mind. Although the current study revealed valuable insights into respondents' LAL needs and levels based on the quantitative and qualitative data, the real classroom assessment practices might yield more reliable and valid data in terms of the active usage of language assessment techniques and methods. In doing so, LAL effect on both students' learning and teachers' teaching practices could be observed.

Moreover limitations of this study need to be considered in the following studies. Therefore, close inspection of LAL development or LAL competences based on the responses of language teachers through qualitative studies from reflective and critical aspects employing observation, reflective journals, think aloud protocols, which may produce in depth understanding of the issue provides fertile ground for further investigations. More importantly, future research could handle the issue across various stakeholders' point of views, namely language learners, native or non native language teachers, teacher trainers, language testers, administrators, policy makers as their perspectives also have an impact on the assessment practices. What's more, for LAL development, the future inquiry might focus on designing a sample professional development program for LAL awareness and development based on the framework designed by the researcher herself for LAL development. Besides, another suggestion for further research involves unveiling how LAL awareness and satisfactory level of LAL facilitate the instructional purposes which in turn enhances student learning.

Considering other variables, the conclusions drawn from the study might prompt future researchers to investigate the effect of certain factors such as gender, teaching experience, educational background, target teaching level of in-service language teachers, school setting and administration, being in a specific language group, etc. on LAL development and assessment beliefs of teachers. Thus, in a future research, it might be useful to formulate a question as what variables would have an impact on LAL of students and teachers to identify the effective factors. For example, the target teaching group such as young learners might change the way of assessment of the teacher. In this sense, a future study might focus on identifying LAL competences or knowledge with a particular attention to language assessment of young learners.

In terms of the timing of the study, as discussed earlier in the document analysis, the language assessment course is offered at 8th semester in spring term. Therefore the current study attempted to explore the LAL needs and levels of pre-service EFL teachers after attending the course at the end of the 8th term. For this reason a future survey could question the awareness of LAL or preparedness level before taking the course as a pre test, and further after taking the course a post test could be administered bearing the timing of the course in mind. Because, the difference between the pre test and post test could imply the impact of the LTA course on language learners' LAL development. Thus timing of the study could change and yield interesting results.

Finally, the research design of this study could be replicated with a satisfactory number of participants. Hopefully, the researchers would conduct a state wide research in order to validate and support the results of this current study in future.

REFERENCES

- AFT, NCME, & NEA. (1990). *Standards for teacher competence in educational assessment of students*. Washington: National Council on Measurement in Education.
- Alam, T. (2017). "English language teaching and learning at secondary and higher secondary levels in Bangladesh". *Journal of Asian and African Social Science and Humanities*, 3 (2), 30-40.
- Alderson, J. C. (2005). *Diagnosing foreign language proficiency: The interface between learning and assessment*. London: Continuum.
- Arslan, R., Bircan, H. and Erođlu, H. (2019). "Cumhuriyet üniversitesi'nde uzaktan eğitimde sunulan derslere yönelik tutum ölçeğinin geliştirilmesi". *Cumhuriyet Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 20(2), 409-427. Retrieved from esjournal.cumhuriyet.edu.tr/tr/pub/issue/50375/530503.
- Anderson, L. W., Krathwohl, D. R. and Bloom, B. S. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives* (Complete ed.). Longman.
- Ashraf, A. and Zaki, S. (2019). "Language assessment in higher education: Challenges and consequences". *Pakistan Journal of Education*, 36(3), 163-185.
- Aydın, M., Baki, A., Köğce, D. and Yıldız, C. (2009). "Mathematics teacher educators' beliefs about assessment". *Procedia Social and Behavioral Sciences*, 1, 2126–2130. <https://doi.org/10.1016/j.sbspro.2009.01.372>
- Babaii, E. and Asadnia, F. (2019). "A long walk to language assessment literacy: EFL teachers' reflection on language assessment research and practice". *Reflective Practice*, 20(6), 745-760. <https://doi.org/10.1080/14623943.2019.1688779>
- Bachman, L. F. (1990). *Fundamental considerations in language testing* . Oxford: Oxford University Press.
- Bachman, L. F. and Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford: Oxford University Press.

- Bailey, K. M., & Brown, J. D. (1995). "Language testing courses: What are they?". A. Cumming and R. Berwick (eds.) in *Validation in language testing* , (pp. 236-256). Cleve-ton, UK: Multilingual Matters.
- Bakanlığı, T. S. (2020). *Haberler*. Retrieved from <https://www.saglik.gov.tr-/Genel/MansetHaberListesi.aspx>
- Baker, B. A. and Riches, C. (2018). "The development of EFL examinations in Haiti: Collaboration and language assessment literacy development". *Language Testing*, 35 (4), 557-581. doi.org/10.1177/0265532217716732
- Ballıdağ, S. and Inan Karagül, B. (2021). "Exploring the language assessment literacy of Turkish in-service EFL teachers". *Balikesir Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 24 (45), 73-92. doi.org/10.31795/baunsobed.909953
- Barnes, N., Fives, H. and Dacey, C. M. (2015). "Teachers' beliefs about assessment". H. Fives, and M. G. Gill (eds.) in *The handbook of research on teachers' beliefs* (pp. 229-233). New York, NY: Routledge.
- Barnes, N., Fives, H. and Dacey, C. M. (2017). "U.S. teachers' conceptions of the purposes of assessment". *Teaching and Teacher Education*, 65, 107-116. doi.org/10.1016/j.tate.2017.02.017
- Brindley, G. (2001). "Outcomes based assessment in practice: Some examples and emerging insights". *Language Testing*, 18(4), 393-407. doi.org/10.1177/026553220101800405
- Brookhart, S. M. (2011). "Educational assessment knowledge and skills for teachers". *Educational Measurement: Issues and Practice*, 30(1), 3-12. doi.org/10.1111/j.1745-3992.2010.00195.x
- Brown, G. T. L. (2006). "Teachers' conceptions of assessment: Validation of an abridged instrument". *Psychological Reports*, 99(1), 166-170. doi.org/10.2466/pr0.99.1.166-170
- Brown, G. T. L. (2008). *Conceptions of assessment: Understanding what assessment means to teachers and students*. New York, NY: Nova Science Publishers.

- Brown, G. T. L., Chaudhry, H. and Dhamija, R. (2015). "The impact of an assessment policy upon teachers' self-reported assessment beliefs and practices: A quasi experimental study of Indian teachers in private schools". *International Journal of Educational Research*, 71, 50-64. doi.org/10.1016/j.ijer.2015.03.001
- Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. Turkey: Pearson Education Inc.
- Brown, J. D. and Bailey, K. M. (2008). "Language testing courses: What are they in 2007?" *Language Testing*, 25(3), 349-383. doi.org/10.1177/0265532208090157
- Brown, J. D. and Hudson, T. (1998). "The alternatives in language assessment". *TESOL Quarterly*, 32(4), 653-675. doi.org/10.2307/3587999
- Bryman, A. (2015). *Social research methods*. Oxford University Press.
- Butler, Y. G. (2022). "Exploring the idea of "Assessment with Learners" in discussing language assessment literacy". *Language Teaching Research Quarterly*, 29, 92-101. doi.org/10.32038/ltrq.2022.29.06
- Butler, Y. G., Peng, X. and Lee, J. (2021). "Young learners' voices: Towards a learner-centered approach to understanding language assessment literacy". *Language Testing*, 38 (3), 429-455. doi.org/10.1177/0265532221992274
- Büyükkarıcı, K. (2014). "Assessment beliefs and practices of language teachers in primary education". *International Journal of Instruction*, 7(1), 1308-1470.
- Büyükkarıcı, K. (2016). "Identifying the areas for English language teacher development: A study of assessment literacy". *Pegem Eğitim ve Öğretim Dergisi*, 6(3), 333-346.
- Canale, M. and Swain, M. (1980). "Theoretical bases of communicative approaches to second language teaching and testing". *Applied Linguistics*, 1, 1-47.
- Carroll, J. D. (1961). *Fundamental considerations in testing for English language proficiency of foreign students*. *Testing*. Washington: Center for Applied Linguistics.
- Chan, C. K. and Jiahui, L. (2020). "A four-dimensional conceptual framework for student assessment literacy in holistic competency development". *Assessment & Evaluation in Higher Education*, 46 (3), 451-466. doi.org/10.1080/02602938.2020.1777388

- Chang, C. W. (2006). "Teachers' beliefs towards oral language assessment in Taiwan collegiate EFL classrooms". *International Conference on English Instruction and Assessment*.
- Charmaz, K. (2014). *Constructing grounded theory*. Los Angeles: Sage Publications.
- Chen, J. C., Dobinson, T. and Kent, S. (2020). "Students' perspectives on the impact of Blackboard Collaborate on Open University Australia (OUA) online learning". *Journal of Educators Online*, 17(1).
- Cheng, L., Rogers, T. and Hu, H. (2004). "ESL/EFL instructors' classroom assessment practices: purposes, methods, and procedures". *Language Testing*, 21(3), 360–389. doi.org/10.1191/0265532204lt288oa
- Choi, I. (2008). "The impact of EFL testing on EFL education in Korea". *Language Testing*, 25(1), 39-62. DOI: 10.1177/0265532207083744
- Chung, S. J. and Nam, Y. (2018). "Language assessment literacy of Korean EFL teachers: An investigation of their training experiences and needs". *Modern English Education*, 19(1), 38-48.
- Cicillini, S. and Giacosa, A. (2020). "Online English-medium instruction (EMI) classes. What we have learned so far". *Conference Paper*, (pp. 178-185).
- COHE. (2018). *Yeni öğretmen yetiştirme lisans programları*. Retrieved from <https://www.yok.gov.tr/kurumsal/idari-birimler/egitim-ogretim-dairesi/yeni-ogretmen-yetistirme-lisans-programlari>
- Cohen, L., Manion, L. and Morrison, K. (2007). *Research methods in education* (Sixth ed.). London and New York: Routledge.
- Compton, L. (2009). "Preparing language teachers to teach language online: A look at skills, roles, and responsibilities". *Computer Assisted Language Learning*, 22 (1), 73-99. doi.org/10.1080/09588220802613831
- Cook, T. D. and Campbell, D. T. (1979). *Quasi-Experimentation: Design and analysis issues for field settings*. Houghton Mifflin.

- Coombe, C., Vafadar, H. and Mohebbi, H. (2020). "Language assessment literacy: what do we need to learn, unlearn, and relearn?" *Language Testing in Asia*, 10(3), 1-16. doi.org/10.1186/s40468-020-00101-6
- Cosgun-Ogeyik, M. (2009). "Evaluation of English language teaching education curriculum by student teachers". *Insan ve Toplum*, 9 (1). 42-57.
- Council, B. (2013). *Turkey national needs assessment of state school English language teaching*. Ankara.
- Creswell, J. W. (2007). *Qualitative inquiry & research design: choosing among five approaches*. Thousand Oaks, CA: Sage Publications.
- Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Sage Publications.
- Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Thousand Oaks, CA: Sage Publications.
- Creswell, J. and Poth, C. (2018). *Qualitative inquiry and research design choosing among five approaches* (4th ed.). Thousand Oaks.: SAGE Publications Inc.
- Crusan, D., Plakans, L. and Gebril, A. (2016). "Writing assessment literacy: Surveying second language teachers' knowledge, beliefs, and practices". *Assessing Writing*, 28, 43-56. doi.org/10.1016/j.asw.2016.03.001
- Çakır, O., Karademir, T. and Erdogdu, F. (2018). "Psychological variables of estimating distance learners' motivation". *Turkish Online Journal of Distance Education*, 19 (1), 163-182.
- Çankaya, P. (2015). "Evaluation of the primary 3rd grade English language teaching program through teachers' eyes". *ELT Research Journal*, 4 (3), 227-238.
- Davidson, P. and Coombe, C. (2019). "Language assessment literacy in the Middle East and North Africa (MENA) Region". *Arab Journal of Applied Linguistics*, 4 (2), 1-23.
- Davies, A. (2008). "Textbook trends in teaching language testing". *Language Testing*, 25(3), 327-347. doi.org/10.1177/0265532208090156

- Delandshere, G. and Jones, J. H. (1999). "Elementary teachers' beliefs about assessment in mathematics: A case of assessment paralysis". *Journal of Curriculum and Supervision*, 14 (3), 216–240.
- DeLuca, C. (2012). "Preparing teachers for the age of accountability: Toward a framework for assessment education". *Action in Teacher Education*, 34 (5-6), 576-591. doi.org/10.1080/01626620.2012.730347
- DeLuca, C. and Klinger, D. A. (2010). "Assessment literacy development: Identifying gaps in teacher candidates' learning". *Assessment in Education: Principles, Policy and Practice*, 17(4), 419–438. doi.org/10.1080/0969594X.2010.516643
- DeLuca, C., Klinger, D., Pyper, J. and Woods, J. (2015). "Instructional rounds as a professional learning model for systemic implementation of assessment for learning". *Assessment in Education: Principles, Policy and Practice*, 22(1), 122-139. doi.org/10.1080/0969594X.2014.967168
- Dörnyei, Z. (2003). *Questionnaires in second language research: Construction, administration, and processing*. Mahwah, Newyork, London: Lawrence Erlbaum Associations.
- Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative and mixed methodologies*. Oxford: Oxford University Press.
- Dörnyei, Z. and Taguchi, T. (2010). *Questionnaires in second language research: Construction, administration and processing* (2nd ed.). New York: Routledge.
- EAPRO. (2020). *Guidance: Assessing and monitoring learning during the Covid-19 crisis*. elttguide.com. (2020). *Top 5 characteristics of a good language test*. Retrieved from elttguide.com.
- Elzainy, A., El Sadik, A. and Al Abdulmonem, W. (2020). "Experience of e-learning and online assessment during the COVID-19 pandemic at the College of Medicine, Qassim University". *J Taibah Univ Med Sc*, 15(6), 456-462. doi.org/10.1016/j.jtumed.2020.09.005
- Farhady, H. (1979). "The disjunctive fallacy between discrete-point and integrative tests". *TESOL Quarterly*, 13(3). doi.org/10.2307/3585882

- Filer, A. (2000). *Assessment: Social practice and social product*. London: Routledge Falmer.
- Freeman, D. and Johnson, K. E. (1998). "Reconceptualizing the knowledge-base of language teacher education". *TESOL Quarterly*, 32(3), 397–417. doi.org/10.2307/3588114
- Freire, P. (1970). *Pedagogy of the oppressed*. New York, NY: Herder & Herder.
- Fulcher, G. (2010). *Practical language testing*. Routledge.
- Fulcher, G. (2012). "Assessment literacy for the language classroom". *Language Assessment Quarterly*, 9(2), 113-132. doi.org/10.1080.15434303.2011.642041
- Fulcher, G. (2020). "Operationalising assessment literacy". D. Tsagari (Ed.) in *Assessment Literacy*. Cambridge Scholars Publishing.
- Fulcher, G., Panahi, A. and Mohebbi, H. (2022). "Glenn Fulcher's thirty-five years contribution to language testing and assessment: A systematic review". *Language Teaching Research Quarterly*, 29, 20-56. doi.org/10.32038/ltrq.2022.29.03
- Gardner, S. and Rea-Dickins, P. (2001). "Conglomeration or chameleon? Teachers representations of language in the assessment of learners with English as an additional language". *Language Awareness*, 10, 161-177. doi.org/10.1080.09658410108667032
- Giraldo, F. (2018). "Language assessment literacy: Implications for language teachers". *Profile: Issues in Teachers' Professional Development*. 20(1), 179-195. doi.org/10.15446/profile.v20n1.62089
- Giraldo, F. (2019). "Language assessment practices and beliefs: Implications for language assessment literacy". *How Journal*, 26(1), 35-61. doi.org/19183/how.26.1.481
- Giraldo, F. (2021). "Language assessment literacy: Insights for educating English language teachers through assessment". *HOW Journal*, 28 (3), 78-92. doi.org/10.19183/how.28.3.673
- Giraldo, F. and Murcia, D. (2018). "Language assessment literacy for pre-service teachers: Course expectations from different stakeholders". *GIST Education and Research Learning Journal*, 16, 56-77.

- Giri, R. A. (2003). "Language testing: Then and now". *Journal of NELTA*, 8(I-2), 49-67. doi.10.3126/nelta.v8i1.3379
- Gürsoy, E., Korkmaz, Ş. Ç. and Damar, E. A. (2013). "Foreign language teaching within 4+4+4 education system in Turkey: Language teachers' voices". *Eurasian Journal of Educational Research*, 53 (A), 59-74.
- Hakim, B. (2015). "English language teachers' ideology of ELT assessment literacy". *International Journal of Education & Literacy Studies*, 3(4). doi.org/10.7575/aiac.ijels.v.3n.4p.42
- Hampel, R. and Stickler, U. (2005). "New skills for new classrooms: Training tutors to teach languages online". *Computer Assisted Language Learning*, 18(4), 311-326.
- Hamp-Lyons, L. (2016). "Purposes of assessment". Banerjee and Tsagari (eds.) in *Handbook of Second Language Assessment*.
- Han, T. and 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. doi.org/10.5296/jse.v4i1.4873
- Harding, L. and Kremmel, B. (2016). "Teacher assessment literacy and professional development". D. Tsagari, and J. Banerjee (eds.) in *Handbook of second language assessment* (pp. 413-428). De Gruyter, Germany.
- Hasselgreen, A., Carlsen, C. and Helness, H. (2004). *European survey of language testing and assessment needs. Part 1: General findings*. Gothenburg, Sweden: European Association for Language Testing and Assessment.
- Hatipoğlu, Ç. (2010). "Summative evaluation of an English language testing and evaluation course for future English language teachers in Turkey". *ELTED*, 13, 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. <https://dergipark.org.tr/en/pub/eltrj/issue/28780/308006>
- Hatipoğlu, Ç. (2017). "History of English language teacher training and English language testing and evaluation (ELTE) education in Turkey". Y. Bayyurt, & N. S. Sifakis

- (eds.) in *English Language Education Policies and Practices in the Mediterranean Countries and Beyond* (pp. 227-257).
- Herrera, L. and Macias, D. (2015). "A call for language assessment literacy in the education and development of teachers of English as a foreign language". *Colomb. Appl. Linguist. J*, 17 (2), 302-312. doi.org/10.14483/udistrital.jour.calj.2015.2.a09
- Hill, M. F., Ell, F., Grudnoff, L. and Limbrick, L. (2014). "Practise what you preach: Initial teacher education students learning about assessment". *Assessment Matters*, 7, 90-112. doi:10.18296/am.0126
- Hymes, D. (1972). "On communicative competence". J. B. Pride and J. Holmes (eds.) in *Sociolinguistics*. Harmondsworth, England: Penguin Books.
- Inbar-Lourie, O. (2008). "Constructing a language assessment knowledge base: A focus on language assessment courses". *Language Testing*, 25(3), 385-402. doi.org/10.1177/0265532208090158
- Inbar-Lourie, O. (2013). "Language assessment literacy: What are the ingredients?" *4th CBLA SIG Symposium Programme*. University of Cyprus.
- Inbar-Lourie, O. (2017). "Language assessment literacy". Shomamy, E., Or, I., May, S. (eds.) in *Language Testing and Assessment, Encyclopedia of Language and Education*. Springer International Publishing. doi.org/10.1007/978-3-319-02261-1_19
- Irez, S. (2007). "Reflection-oriented qualitative approach in beliefs research". *Eurasia Journal of Mathematics, Science & Technology Education*, 3 (1), 17-27. doi.org/10.12973/ejmste/75370
- Jeong, H. (2013). "Defining assessment literacy: Is it different for language testers and non-language testers?" *Language Testing*, 30 (3), 345-362. doi.org/10.1177/0265532213480334
- 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. doi.org/10.1177/0265532209351431

- Karataş, T. Ö. and Tuncer, H. (2020). "Sustaining language skills development of pre-service EFL teachers despite the COVID-19 interruption: A case of emergency distance education". *Sustainability*, 12, 8188. doi.org/10.3390/su12198188
- Kavaklı, N. and Arslan, S. (2019). "Towards a continuum from know-how to show-how for developing EFL student-teachers' assessment literacy". *International Online Journal of Education and Teaching (IOJET)*, 6 (1). 223-232.
- Kaya, T. and Mede, E. (2021). "Exploring language assessment literacy of EFL instructors in language preparatory programs". *Eğitim Fakültesi Dergisi*, 7 (1), 163-189. DOI: 10.17932/IAU.EFD.2015.013/efd_v07i008
- Kırkgöz, Y. (2007). "English language teaching in Turkey: Policy changes and their implementations". *SAGE Publications*, 38 (2), 216-228. doi.org/10.1177/0033688207079696
- Kırkgöz, Y. (2008). "Curriculum innovation in Turkish primary education". *Asia-Pacific Journal of Teacher Education*, 36 (4), 309-322. doi.org/10.1080/13598660802376204
- Kırkgöz, Y. (2009). "Globalization and English language policy in Turkey". *Educational Policy*, 23 (5), 663-684. doi.org/10.1177/0895904808316319
- Koh, K. H. (2011). "Improving teachers' assessment literacy through professional development". *Teaching Education*, 22 (3), 255-276. doi.org/10.1080/10476210.2011.593164
- Koh, K., Burke, L. E., Luke, A., Gong, W. and Tan, C. (2018). "Developing the assessment literacy of teachers in Chinese language classrooms: A focus on assessment task design". *Language Teaching Research*, 22(3), 264-288. doi.org/10.1177/1362168816684366
- Komur, S. (2018). "Preservice English teachers' assessment awareness: Level of readiness for classroom practice". *The Journal of Language Teaching and Learning*, 8 (1), 109-121. Retrieved from <https://www.jltl.com.tr/index.php/jltl/article/view/60>
- Köksal, D. (2004). "Assesing teachers' testing skills in ELT and enhancing their professional development through distance learning on the net". *Turkish Online*

Journal of Distance Education, 5 (1), 1-11. Retrieved from <https://dergipark.org.tr/tr/pub/tojde/issue/16931/176755>

- König, J., Jäger-Biela, D. J. and Glutsch, N. (2020). “Adapting to online teaching during COVID-19 school closure: Teacher education and teacher competence effects among early career teachers in Germany”. *European Journal of Teacher Education*, 43(4), 608-622. doi.org/10.1080/02619768.2020.1809650
- Kremmel, B. and Harding, L. (2020). “Towards a comprehensive, empirical model of language assessment literacy across stakeholder groups: Developing the language assessment literacy survey”. *Language Assessment Quarterly*, 17(1), 100–120. doi.org/10.1080/15434303.2019.1674855
- Krueger, K. (2015). *How to catch students cheating on online tests*. Retrieved from <http://mediashift.org/2015/08/how-to-catch-students-cheating-on-online-tests/>
- Kruse, L., Impellizzeri, W., Witherel, C. E. and Sondergeld, T. A. (2020). “Evaluating the impact of an assessment course on preservice teachers’ classroom assessment literacy and self-efficacy”. *Mid-Western Educational Researcher*, 32(2), 107-132.
- Lam, R. (2015). “Language assessment training in Hong Kong: Implications for language assessment literacy”. *Language Testing*, 32 (2), 169-197. doi.org/10.1177/0265532214554321
- Lam, R. (2019). “Teacher assessment literacy: Surveying knowledge, conceptions and practices of classroom-based writing assessment in Hong Kong”. *System*, 81, 78–89. doi.org/10.1016/j.system.2019.01.006
- Levi, T. and Inbar-Lourie, O. (2020). “Assessment literacy or language assessment literacy: Learning from the teachers”. *Language Assessment Quarterly*, 17 (2), 168-182. doi.org/10.1080/15434303.2019.1692347
- Li, H., Zhong, Q. and Suen, H. K. (2012). “Students’ perceptions of the impact of the college English test”. *Language Testing in Asia*, 2(3), 77-94. doi.org/10.1186/2229-0443-2-3-77
- Lie, A., Tamah, S. M., Gozali, I., Triwidayati, K. R., Utami, T. S. and Jemadi, F. (2020). “Secondary school language teachers’ online learning engagement during the

- Covid-19 pandemic in Indonesia”. *Journal of Information Technology Education: Research*, 19, 803-832. doi.org/10.28945/4626
- Lincoln, Y. S. and Guba, E. G. (1985). *Naturalistic inquiry*. Newbury Park, CA: Sage Publications.
- Lynch, B. K. (1996). *Language program evaluation: Theory and practice*. New York: Cambridge University Press.
- Mackey, A. and Gass, S. M. (2005). *Second language research: Methodology and design*. Lawrence Erlbaum Associates Publishers.
- Mahapatra, S. (2019). *Evolution of English language assessment: A critical account*. Retrieved from <https://theonlyeducationist.blogspot.com/2019/11/evolution-of-english-language.html>
- Malone, M. (2013). “The essentials of assessment literacy: Contrasts between testers and users”. *Language Testing*, 30 (3), 329–344. doi.org/10.1177/0265532213480129
- Matese, G., Griesdorn, J. and Edelson, D. C. (2002). Teacher beliefs about assessment in inquiry-based science: Research to inform professional development. In annual meeting of *American Educational Research Association Conference*.
- Maxwell, J. A. (2010). “Using Numbers in Qualitative Research”. *Qualitative Inquiry*, 16(6), 475–482. DOI: 10.1177/1077800410364740
- McCallum, L. and Coombe, C. (2022). “An interview with Glenn Fulcher: Looking back and moving forward”. *Language Teaching Research Quarterly*, 29, 6-19. doi:10.32038/ltrq.2022.29.02
- McDonough, J. and McDonough, S. (2006). *Research Methods for English Language Teachers*. London: Arnold.
- Mede, E. and Atay, D. (2017). “English language teachers’ assessment literacy: The Turkish context”. *Ankara Üniversitesi TÖMER Dil Dergisi*, 168(1), 43-60.
- Mertler, C. A. (2003). Pre-service versus in-service teachers’ assessment literacy: Does classroom experience make a difference? *Annual meeting of the Mid-Western Educational Research Association*. Columbus, OH.

- Mertler, C. A. (2009). "Teachers' assessment knowledge and their perceptions of the impact of classroom assessment professional development". *Improving Schools*, 12 (2), 101–113. doi.org/10.1177/1365480209105575
- Miles, M. B. and Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. (2nd ed). Thousand Oaks, CA: Sage.
- Miller, M. D., Linn, R. L. and Gronlund, N. (2013). *Measurement and assessment in teaching* (11th ed.). New York: Pearson Education.
- Mohajer, Z. (2020). "Native and non-native English teachers' perceptions toward assessment literacy regarding their teaching experience". *International Journal of Latest Research in Humanities and Social Science (IJLRHSS)*, 3(1), 58-65.
- Muñoz, A. P., Palacio, M. and Escobar, L. (2012). "Teachers' beliefs about assessment in an EFL context in Colombia". *PROFILE*, 14(1), 143-158.
- O'Loughlin, K. (2006). "Learning about second language assessment: Insights from a postgraduate student on-line forum". *University of Sydney Papers in TESOL*, 1 (1), 71–85.
- Ogan-Bekiroglu, F. and Suzuk, E. (2014). "Pre-service teachers' assessment literacy and its implementation into practice". *Curriculum Journal*, 25 (3), 344-371. doi.org/10.1080/09585176.2014.899916
- Opre, D. (2015). "Teachers' conceptions of assessment". *Procedia - Social and Behavioral Sciences*, 209, 229 – 233. doi.org/10.1016/j.sbspro.2015.11.222
- Ölmezer-Öztürk, E. and Aydın, B. (2018). *Developing and validating language assessment knowledge scale (LAKS) and exploring the assessment knowledge of EFL teachers*. Unpublished PhD Dissertation, Anadolu University.
- Öz, S. and Atay, D. (2017). "Turkish EFL instructors' in-class language assessment literacy: Perceptions and practices". *ELT Research Journal*, 6 (1), 25-44. Retrieved from <https://dergipark.org.tr/tr/pub/eltrj/issue/28779/307996>.
- Pajares, M. F. (1992). "Teachers' beliefs and educational research: Cleaning up a messy construct". *Review of Educational Research*, 62 (3), 307-332. doi.org/10.3102/00346543062003307

- Paradowski, M. B. (2002). "The features of good language tests". *The Teacher*, 1 (40). Retrieved from <https://files.eric.ed.gov/fulltext/ED503442.pdf>
- Popham, W. J. (2014). *Classroom assessment: What teachers need to know* (7th ed.). Boston: Pearson Education.
- Rahim, A. F. (2020). "Guidelines for online assessment in emergency remote teaching during the COVID-19 pandemic". *Education in Medical Journal*, 12 (2), 59–68. doi.org/10.21315/eimj2020.12.2.6
- Reynolds-Keefers, L. (2010). "Rubric-referenced assessment in teacher preparation: An opportunity to learn by using". *Practical Assessment, Research and Evaluation*, 15 (8), 1-9. doi.org/10.7275/psk5-mf68
- Rogers, W. T., Cheng, L. and Hu, H. (2007). "ESL/EFL Instructors' beliefs about assessment and evaluation". *Canadian and International Education / Education*, 36 (1).
- Rueda, R. and Garcia, E. (1996). "Teachers' perspectives on literacy assessment and instruction with language-minority students: A comparative study". *The Elementary School Journal*, 96 (3), 311-332. Retrieved from <http://www.jstor.org/stable/1001760>
- Russell, M. K. and Airasian, P. W. (2012). *Classroom assessment: Concepts and applications* (7th ed.). New York, NY: McGraw-Hill.
- Saputra, E., Hamied, F. and Suherdi, D. (2020). "The development of beliefs and practices of language assesment literacy: Does a professional learning community help?" *Journal of Education for Teaching*, 46 (3), 414-416. doi.org/10.1080/02607476.2020.1761250
- 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.org/10.1177/0265532213480128
- Semiz, Ö. and Odabaş, K. (2016). Turkish EFL teachers' familiarity of and perceived needs for language testing and assessment literacy. *3rd international linguistics and language studies conference*. İstanbul, Turkey.

- Sevimel Şahin, A. (2019). *Exploring foreign language assessment literacy of pre-service English language teachers*. Unpublished Doctoral Dissertation, Anadolu University, Eskişehir, Turkey.
- Shenton, A. K. (2004). "Strategies for ensuring trustworthiness in qualitative research". *Education for Information*, 22, 63-75. doi.org/10.3233/EFI-2004-22201
- Shohamy, E. (2001). "Democratic assessment as an alternative". *Language Testing*, 18 (4), 373 – 392.
- Skott, J. (2015). "The promises, problems, and prospects of research on teachers' beliefs". Fives, H. and Gill M. G (eds.) in *International handbook of research on teachers' beliefs* (pp. 13-31). New York and London: Routledge Taylor & Francis Group.
- Smedley, J. (2010). "Modelling the impact of knowledge management using technology". *OR Insight*, 23 (4), 233-250. doi.org/10.1057/ori.2010.11
- Spiller, D. (2012). "Assessment matters: Self-assessment and peer assessment". *Teaching Development*. University of Waikato.
- Spolsky, B. (1978). *Approaches to language testing*. Advances in language testing series: 2. Papers in Applied Linguistics.
- Stiggins, R. J. (1999). "Evaluating classroom assessment training in teacher education". *Educational Measurement: Issues and Practice*, 18 (1), 23-27. doi.org/10.1111/j.1745-3992.1999.tb00004.x
- Strauss, A. L. (1987). *Qualitative analysis for social scientists*. Cambridge University Press.
- Sultana, N. (2019). "Language assessment literacy: An uncharted area for the English language teachers in Bangladesh". *Language Testing in Asia*, 9 (1), 2-14. doi.org/10.1186/s40468-019-0077-8
- Şahin, S. (2018). *An analysis of English language testing and evaluation course in English language teacher education programs in Turkey: developing language assessment literacy of pre-service EFL teachers*. Unpublished doctoral dissertation, Middle East Technical University.

- Şahinkarakas, Ş. (2012). "The role of teaching experience on teachers' perceptions of language assessment". *Procedia - Social and Behavioral Sciences*, 47. doi.org/10.1016/j.sbspro.2012.06.901
- Sendoğan, K. (2020). "Foreign language education during Covid-19 pandemic: An evaluation from the perspectives of preparatory class students". *MİLLÎ EĞİTİM*, 49 (1), 1079-1090. doi.org/10.37669/milliegitim.788274
- Şişman, P. E. and Büyükkarcı, K. (2019). "A review of foreign language teachers' assessment literacy". *Sakarya University Journal of Education*, 9 (3), 628-650. doi.org/10.19126/suje.621319
- Tao, N. (2014). *Development and validation of classroom assessment literacy scales: English as a foreign language (EFL) instructors in a Cambodian higher education setting*. Unpublished PhD Dissertation, Victoria University, Melbourne, Australia.
- Taopan, L. L., Drajati, N. A. and Sumardi (2020). "TPACK framework: Challenges and opportunities in EFL classrooms". *Research and Innovation in Language Learning*, 3 (1), 1-22. doi. 10.33603/rill.v3i1.2763
- Taylor, L. (2009). "Developing assessment literacy". *Annual Review of Applied Linguistics*, 21-36. doi.org/10.1017/S0267190509090035
- Taylor, L. (2013). "Communicating the theory, practice and principles of language testing to test stakeholders: Some reflections". *Language Testing*, 30 (3), 403-412. doi.org/10.1177/0265532213480338
- Thomas, M. (2012). "Teachers' beliefs about classroom assessment and their selection of classroom assessment strategies". *Journal of Research and Reflections in Education*, 6 (2), 103 -112. Retrieved from <http://www.ue.edu.pk/journal.asp>
- Topkaya, E. Z. and Küçük, Ö. (2010). "An evaluation of 4th and 5th grade English language teaching program". *İlkogretim Online*, 9 (1), 52-65. Retrieved from <https://dergipark.org.tr/tr/pub/ilkonline/issue/8596/106953>
- Tsagari, D. (2016). "Assessment orientations of state primary EFL teachers in two mediterranean countries". *CEPS Journal*, 6 (1), 9-30. doi.org/10.26529/cepsj.102

- Tsagari, D. and 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. Retrieved from <http://ro.ecu.edu.au/ajte/vol37/iss7/2>
- Uztosun, M. S. (2018). "In-service teachers' voices on the quality of English language teaching in Turkey". *Erzincan Üniversitesi Eğitim Fakültesi Dergisi*, 20 (1), 140-155. DOI: [10.17556/erziefd.334460](https://doi.org/10.17556/erziefd.334460)
- Uztosun, M. S. (2018a). "Professional competences to teach English at primary schools in Turkey: a Delphi study". *European Journal of Teacher Education*, 41 (4), 549-565. doi.org/10.1080/02619768.2018.1472569
- Vogt, K. and Tsagari, D. (2014). "Assessment literacy of foreign language teachers: Findings of a European study". *Language Assessment Quarterly*, 11(4), 374-402. doi.org/10.1080/15434303.2014.960046
- Volante, L. and Fazio, X. (2007). "Exploring teacher candidates' assessment literacy: Implication for teacher education reform and professional development". *Canadian Journal of Education*, 30(3), 749. Retrieved from <https://journals.sfu.ca/cje/index.php/cje-rce/article/view/2973>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Weber, R. (1990). *Basic Content Analysis*. London: Sage Publications.
- Weideman, A. (2019). "Assessment literacy and the good language teacher: four principles and their applications". *Journal for Language Teaching*, 53(1), 103–121. DOI:10.4314/jlt.v53i1.5
- Wicking, P. (2017). "The assessment beliefs and practices of English teachers in Japanese universities". *JLTA Journal*, 20, 76–89. doi.org/10.20622/jltajournal.20.0_76
- Williams, C. H. (2017). *Teachers in East Asia: A teacher's guide to Chinese, Japanese, and Korean learners*. Singapore: Springer.

- Winke, P. (2011). "Evaluating the validity of a high-stakes ESL test: Why teachers' perceptions matter". *TESOL Quarterly*, 45(4), 628-660. doi.org/10.5054/tq.2011.268063
- Xu, Y. and Brown, G. T. (2016). "Teacher assessment literacy in practice: A reconceptualization". *Teaching and Teacher Education*, 58, 149-162. doi.org/10.1016/j.tate.2016.05.010
- Yan, X. and Fan, J. (2020). "Am I qualified to be a language tester?": Understanding the development of language assessment literacy across three stakeholder groups". *Language Testing*, 0 (0), 1-28. doi.org/10.1177/0265532220929924
- Yan, X., Zhang, C. and Fan, J. J. (2018). "Assessment knowledge is important, but...: How contextual and experiential factors mediate assessment practice and training needs of language teachers". *System*, 74, 158–168. DOI: [10.1016/j.system.2018.03.003](https://doi.org/10.1016/j.system.2018.03.003)
- Yastıbaş, A. E. (2018). *Language assessment literacy of Turkish EFL instructors: a multiple-case study*. Unpublished doctoral dissertation, Atatürk University, Erzurum, Turkey.
- Yastıbaş, A. E. and Takkaç, M. (2018). "Understanding language assessment literacy: Developing language assessment". *Journal of Language and Linguistic Studies*, 14 (1), 178-193.
- Yetkin, C. (2015). *An investigation on ELT teacher candidates' assessment literacy*. Unpublished master's thesis, Çağ University, Mersin, Turkey.
- YÖK. (2020). *Üniversitelerde uygulanacak uzaktan eğitime ilişkin açıklama*. Retrieved from Haberler: Retrieved from <https://www.yok.gov.tr/Sayfalar/Haberler/2020/universitelerde-uygulanacak-uzaktan-egitime-iliskin-aciklama.aspx>

APPENDICES
APPENDIX 1:
CONSENT FORM

English Language Teachers' Assessment Literacy: Identifying Assessment Needs and Beliefs of EFL Teachers and Teacher Candidates

You are being asked to participate in a research study. Before you give your consent to volunteer, it is important that you read the following information and ask as many questions as necessary to be sure you understand what you will be asked to do.

Researcher

This research study is being conducted by Pınar Çankaya, a Ph.D. candidate at ÇOMÜ. The name of the project is “English Language Teachers’ Assessment Literacy: Identifying Assessment Needs and Beliefs of EFL Teachers and Teacher Candidates

Purpose of the research

This research study is designed to investigate the language assessment literacy levels and needs of participants along with their assessment beliefs. In doing so, the researcher attempts to gain insights to give recommendations for improving the language and assessment course offered during pre service English language teacher education.

The procedure

If you volunteer to participate in this study, you will be asked to complete the survey and answer interview questions as well. Your participation will take approximately fifteen minutes for the survey, and 45 minutes for the interview session. The names of all participants will remain confidential. Pseudonyms will be used for the participants when necessary.

Pınar Çankaya

	Please Put a Tick
I confirm that I have read and understand the information for the study and have had the opportunity to ask questions.	<input type="checkbox"/>
I understand that my participation is voluntary and that I am free to withdraw at any time, without giving reason.	<input type="checkbox"/>
I agree to take part in the above study.	<input type="checkbox"/>
I agree to the face to face / focus groups / e-mail interview being audio or video recorded.	<input type="checkbox"/>
I agree to the use of anonymised quotes in this study.	<input type="checkbox"/>

Participant Code

Date

Signature

Name of Researcher

Date

Signature

APPENDIX 2:

OPEN ENDED INTERVIEW QUESTIONS (1st Version)

ASSESSMENT BELIEFS

- 1) How do you think assessment improves student learning?
- 2) How do you think assessment improves your teaching?
- 3) Can you name at least three useful assessment types-methods for student learning?
- 4) How do you make use of assessment results for students' improvement?
- 5) What are your reasons for assessing students?
- 6) What do you think about the features of a sound assessment?
- 7) What are the possible reasons for poor assessment?
- 8) What do you think of assessment for / as / of learning? Can you explain giving examples?
- 9) What do you think of students' responsibility in assessment process?
- 10) What does language assessment literacy" suggest to you?
- 11) Of the topics above (Fulcher's scale), which one / ones do you think the most helpful to language teachers?
- 12) Of the topics above (Fulcher's scale), which one / ones do you think the most challenging for language teachers?

NEEDS FOR ASSESSMENT

1. Do you feel competent when assessing students?
2. What do you think about your needs (knowledge, skills, practice) in assessment process?
3. Do you think language assessment course in pre service teacher education has met your needs of assessment?
4. How do you improve language assessment literacy?
5. What are the possible ways for improving language assessment literacy?

ONLINE ASSESSMENT

- 1) What do you think of online language assessment?

- 2) What kind of tasks-tests can be suitable for online language assessment?
- 3) How do you ensure validity and reliability in online language assessment?
- 4) What challenges you have faced, what coping strategies you have used during online language assessment?
- 5) What do you think of the online feedback? Possible ways, challenges, positive sides and negative sides.



APPENDIX 3:
OPEN ENDED INTERVIEW QUESTIONS (2nd Version)

BİLGİLENDİRİLMİŞ GÖNÜLLÜ ONAM METNİ

Sizi Pınar ÇANKAYA tarafından yürütülen “İngilizce Öğretmenlerinin Dil Değerlendirme Okuryazarlığı: EFL öğretmen ve öğretmen adaylarının dil değerlendirme ihtiyaç ve inanışlarını belirleme” başlıklı araştırmaya davet ediyoruz. Bu araştırmanın amacı katılımcıların “Yabancı Dilde Ölçme ve Değerlendirme Okuryazarlığı (LAL)” düzeylerini keşfetmek ve öğretmen Ölçme ve değerlendirme inanışlarının değerlendirme okuryazarlığı üzerindeki etkilerini sorgulamaktır. Araştırmada sizden tahminen 15-20 dakika ayırmanız istenmektedir. Araştırmaya sizin dışınızda tahminen 100 kişi katılacaktır. Bu çalışmaya katılmak tamamen **gönüllülük** esasına dayanmaktadır. Çalışmanın amacına ulaşması için sizden beklenen, bütün soruları eksiksiz, kimsenin baskısı veya telkini altında olmadan, size en uygun gelen cevapları içtenlikle verecek şekilde cevaplamanızdır. Bu formu okuyup onaylamanız, araştırmaya katılmayı kabul ettiğiniz anlamına gelecektir. Ancak, çalışmaya katılmama veya katıldıktan sonra herhangi bir anda çalışmayı bırakma hakkına da sahipsiniz. Bu çalışmadan elde edilecek bilgiler tamamen araştırma amacı ile kullanılacaktır.

-
- Araştırmaya katılmayı kabul ediyorum.
- Araştırmaya katılmayı kabul etmiyorum.

ASSESSMENT BELIEFS

- 1) What are your reasons for assessing students?
- 2) How do you think assessment affect student learning?
- 3) How do you think assessment affect your teaching?
- 4) Can you name at least three useful assessment types-methods for student learning?
- 5) How do you make use of assessment results for students' improvement?
- 6) What do you think about the features of a good/ sound assessment?
- 7) What are the possible reasons for poor assessment?
- 8) What do you think of assessment for / as / of learning? Can you explain giving examples?
- 9) What do you think of students' responsibility (or role) in assessment process?
- 10) What does “language assessment literacy” mean for you?

- 11) Of the topics above (Fulcher's scale), which one / ones do you think the most helpful to language teachers?
- 12) Of the topics above (Fulcher's scale), which one / ones do you think the most challenging for language teachers?

NEEDS FOR ASSESSMENT

1. Do you feel competent when assessing students? If yes, How? If not, why not?
2. What do you think about your needs (knowledge, skills, practice) in assessment process?
3. Do you think language assessment course in pre service teacher education has met your needs of assessment? If yes, How? If not, why not?
4. How do you improve language assessment literacy?
5. What are the possible ways for improving language assessment literacy?

ONLINE ASSESSMENT

- 1) What do you think of online language assessment?
- 2) What kind of tasks-tests can be suitable for online language assessment?
- 3) How do you ensure validity and reliability in online language assessment?
- 4) What challenges have you faced, what coping strategies have you used during online language assessment?
- 5) What do you think of the online feedback? Possible ways, challenges, positive sides and negative sides.

APPENDIX 4:

LANGUAGE ASSESSMENT LITERACY INVENTORY FOR TEACHERS

Dear participants,

You are invited to fill out a survey that aims to identify your perceptions, beliefs and needs regarding your language assessment literacy. Your responses are very important in order to design a language assessment course. Please be kind enough to give truthful and straightforward answers in order to obtain accurate results. It will take approximately 15 minutes to complete the questionnaire.

Thank you very much for your time and support.

Prof. Dr. Dinçay KÖKSAL

Öğr. Gör. Pınar ÇANKAYA

I voluntarily participate in this study. Yes No

PART 1: LANGUAGE ASSESSMENT LITERACY INVENTORY

Please indicate **your LEVEL** for the following topics with regards to language assessment literacy.

1	2	3	4	5	
Poor	Fair	Average	Good	Excellent	
History of Language Testing	1	2	3	4	5
Procedures in language test design	1	2	3	4	5
Deciding what to test	1	2	3	4	5
Writing test specifications/blueprints	1	2	3	4	5
Writing test tasks and items	1	2	3	4	5
Evaluating language tests	1	2	3	4	5
Interpreting scores	1	2	3	4	5
Test analysis	1	2	3	4	5
Selecting tests for your own use	1	2	3	4	5
Reliability	1	2	3	4	5
Validation	1	2	3	4	5
Use of statistics	1	2	3	4	5
Rating performance tests (speaking/writing)	1	2	3	4	5
Scoring closed-response items	1	2	3	4	5
Classroom assessment	1	2	3	4	5
Large-scale testing	1	2	3	4	5
Standard setting	1	2	3	4	5
Preparing learners to take tests	1	2	3	4	5
Washback on the classroom	1	2	3	4	5
Test administration	1	2	3	4	5
Ethical considerations in testing	1	2	3	4	5
The uses of tests in society	1	2	3	4	5
Principles of educational measurement	1	2	3	4	5

PART 2: LANGUAGE ASSESSMENT LITERACY NEEDS INVENTORY

Please indicate **your NEED** for the following topics with regards to language assessment literacy.

	1	2	3	4	5
	No Need	Low	Moderate	High	Very high
History of Language Testing	1	2	3	4	5
Procedures in language test design	1	2	3	4	5
Deciding what to test	1	2	3	4	5
Writing test specifications/blueprints	1	2	3	4	5
Writing test tasks and items	1	2	3	4	5
Evaluating language tests	1	2	3	4	5
Interpreting scores	1	2	3	4	5
Test analysis	1	2	3	4	5
Selecting tests for your own use	1	2	3	4	5
Reliability	1	2	3	4	5
Validation	1	2	3	4	5
Use of statistics	1	2	3	4	5
Rating performance tests (speaking/writing)	1	2	3	4	5
Scoring closed-response items	1	2	3	4	5
Classroom assessment	1	2	3	4	5
Large-scale testing	1	2	3	4	5
Standard setting	1	2	3	4	5
Preparing learners to take tests	1	2	3	4	5
Washback on the classroom	1	2	3	4	5
Test administration	1	2	3	4	5
Ethical considerations in testing	1	2	3	4	5
The uses of tests in society	1	2	3	4	5
Principles of educational measurement	1	2	3	4	5

PART 3: PERCEIVED IMPORTANCE FOR LANGUAGE ASSESSMENT LITERACY

Please indicate your **PERCEIVED IMPORTANCE** level for the following topics in order to develop language assessment literacy of teachers.

	1	2	3	4	5
	Unimportant	Not very important	Fairly important	Important	Essential
History of Language Testing	1	2	3	4	5
Procedures in language test design	1	2	3	4	5
Deciding what to test	1	2	3	4	5
Writing test specifications/blueprints	1	2	3	4	5
Writing test tasks and items	1	2	3	4	5
Evaluating language tests	1	2	3	4	5
Interpreting scores	1	2	3	4	5
Test analysis	1	2	3	4	5
Selecting tests for your own use	1	2	3	4	5
Reliability	1	2	3	4	5
Validation	1	2	3	4	5
Use of statistics	1	2	3	4	5

Rating performance tests (speaking/writing)	1	2	3	4	5
Scoring closed-response items	1	2	3	4	5
Classroom assessment	1	2	3	4	5
Large-scale testing	1	2	3	4	5
Standard setting	1	2	3	4	5
Preparing learners to take tests	1	2	3	4	5
Washback on the classroom	1	2	3	4	5
Test administration	1	2	3	4	5
Ethical considerations in testing	1	2	3	4	5
The uses of tests in society	1	2	3	4	5
Principles of educational measurement	1	2	3	4	5

PART 5: BACKGROUND INFORMATION

Gender: Female Male **Age:** _____

How long have you been teaching English?: _____ years

Which department did you graduate?

English Language Teaching English Language and Literature Linguistics
 Translating and Interpreting Other: _____

Please indicate your educational background:

Bachelor's degree Master's degree completed or in progress
 PhD degree completed or in progress

Have you ever taken a training /course on language testing and assessment?

No training A little Training Extensive training

Please indicate your PERCEIVED language assessment literacy LEVEL,

POOR FAIR AVERAGE GOOD EXCELLENT

Please indicate your PERCEIVED language assessment literacy NEED,

No need Low need Moderate need High Need Very high need

APPENDIX 5:

A SAMPLE GUIDELINE FOR TEACHING LAL

Principle / Technique: Inappropriateness Analysis	
Description	Analysis of existing sample tests for their appropriateness or inappropriateness in terms of test purpose, test content, test method, test items among others.
Content	LAL Topics such as: writing test items, validity, reliability, test purpose, deciding what to test, selecting tests for their own use, ethical considerations, online language assessment, using statistics, designing alternative assessment tasks, interpreting scores
Materials/Resources	Existing sample language tests / exams developed by teachers or colleagues
Teaching/Learning Strategy	Group discussion, analysing, problem solving, reporting
Activities	Group work, sample test reviews
Roles and Actions	All students are critical and reflective thinkers and evaluators. In each group, some students might perform as note takers and/or reporters.
Learning Outcome	An analysis of a sample language test discussing its weaknesses and strengths by deciding whether the test item/method/purpose is appropriate or not.
Learning Objectives	<ul style="list-style-type: none"> • Understanding what is appropriate and what is not in a language test in terms of test purpose, test content, test method, test items, ethical concerns. • Acquiring the skills and necessary knowledge about test development and test specifications to evaluate a language test and overcoming the inappropriate points by analysing it critically. • Exploring the possible reasons of badly designed tests by determining the source of the problems (e.g. face validity, inappropriate test method, unreliable test scores, among others) and developing coping strategies to solve these problems.
Feedback	Students discuss about existing problems and possible solutions by comparing their reports and analysis notes.
Assignment	Students are expected to write a reflection paper about weaknesses and problems within existing sample tests and possible solutions to design a more valid, reliable and effective test.

APPENDIX 6:

ETİK KURUL



T.C.
ÇANAKKALE ONSEKİZ MART ÜNİVERSİTESİ REKTÖRLÜĞÜ
Lisansüstü Eğitim Enstitüsü Etik Kurulu
Bilimsel Araştırma Etik Kurulu



Sayı : E-84026528-050.01.04-2200064803
Konu : Başvuru İncelenmesi

21.03.2022


Sayın Pınar ÇANKAYA

Yürütücülüğünüzü yapmış olduğunuz 2022-YÖNP-0217 nolu projeniz ile ilgili Bilimsel Araştırmalar Etik Kurulu'nun almış olduğu 17.03.2022 tarih ve 06/07 sayılı kararı aşağıdadır.

Bilgilerinize rica ederim.

KARAR 07- Pınar ÇANKAYA'nın sorumlu yürütücülüğünü yaptığı "English Language Teachers' Assessment Literacy: Identifying Assessment Needs and Beliefs of EFL Teachers, Teacher Candidates, and Teacher Educators" başlıklı araştırmasının, ilgili kurumun izninin alınması ve Bilimsel Araştırmalar Etik Kuruluna sunulması koşulu ile Etik Kurul ilkelerine uygun olduğuna oy birliği ile karar verilmiştir.

APPENDIX 7: PERMISSION FOR DATA TOOLS

PINAR PINAR 
Also seen in
Dear Dr. Fulcher,
My name is Pinar Çankaya. I am an English Language teacher in a state university in Turkey, and a doctoral student at Çanakkale Onsekiz Mart University, as well. For my dissertation I am investigating pre-service and in-service EFL teachers language assessment literacy. I read your paper titled as "Assessment Literacy for the Language Classroom" in which you noted items to be placed in a language testing course. I am kindly asking your permission to use these items in my research.
Thank you in advance for any consideration you would give to discussing the survey with me.
Sincerely,
Pinar ÇANKAYA


Fulcher, Glenn (Prof.) - glenn.fulcher@leicester.ac.uk
Also seen in
12:31 (23 dakika önce) ☆ ↶ ⋮
İngilizce • > Türkçe • İstisnai çeviri İngilizce için kapat x
Dear Pinar,
Your research sounds very interesting! Please do use anything you like from that article.
Best wishes for a successful outcome,
Glenn
Glenn Fulcher PhD
Professor of Applied Linguistics and Language Assessment
glenn.fulcher@leicester.ac.uk
University of Leicester, UK
<http://www.le.ac.uk/~glf20>
ORCID: <https://orcid.org/0000-0001-9000-0001>
W: <http://www.le.ac.uk/~glf20>
Winner of the 2008 SAGE UK Book Award <http://www.sagepub.com/uk/home.html>



APPENDIX 8:
PERMISSION FOR RESEARCH



T.C.
TEKİRDAĞ NAMİK KEMAL ÜNİVERSİTESİ REKTÖRLÜĞÜ
Bilimsel Araştırma ve Yayın Etik Kurulu Başkanlığı

Sayı : E-12394611-044-149926
Konu : Anket İzni - Pınar ÇANKAYA

06.04.2022

DAĞITIM YERLERİNE

İlgi : Öğrenci İşleri Daire Başkanlığının 01.04.2022 tarihli ve 148460 sayılı yazısı

İlgi yazı ile; Çanakkale Onsekiz Mart Üniversitesi Lisansüstü Eğitim Enstitüsü Yabancı Diller Eğitimi Anabilim Dalı İngiliz Dili Eğitimi Doktora Programı öğrencisi Pınar ÇANKAYA'nın, "English Language Teachers' Assessment Literacy:Identifying Assessment Needs and Beliefs of EFL Teachers, Teacher Candidates and Teacher Educators" konulu tez çalışmasını, Üniversitemizde Yabancı Diller Yüksekokullarında ders vermekte olan öğretim görevlilerine uygulama talebine ilişkin olarak alınan yazı bildirilmektedir.

Söz konusu anket iznine ilişkin değerlendirmeler neticesinde anılan talep uygun görülmüş olup, konu hakkında;



T.C.
BURSA TEKNİK ÜNİVERSİTESİ REKTÖRLÜĞÜ
Genel Sekreterlik

Sayı : E-31633970-044-55658
Konu : Anket İzni (Pınar ÇANKAYA)

05.04.2022

ÇANAKKALE ONSEKİZ MART ÜNİVERSİTESİ REKTÖRLÜĞÜNE

İlgi : 01.04.2022 tarih ve E-93130991-302.08.01-2200072658 sayılı yazı.

Üniversiteniz Lisansüstü Eğitim Enstitüsü Yabancı Diller Eğitimi Anabilim Dalı İngiliz Dili Eğitimi Doktora Programı öğrencisi Pınar ÇANKAYA' nın "English Language Teachers' Assessment Literacy: Identifying Assessment Needs and Beliefs of EFL Teachers, Teacher Candidates and Teacher Educators" konulu tez çalışması kapsamında Üniversitemiz Yabancı Diller Yüksekokulunda görev yapan öğretim görevlilerine anket/ölçek araştırması yapmasında araştırmanın etik kurallar çerçevesinde ve gönüllülük esasına dayalı olarak yapılması koşuluyla herhangi bir sakınca bulunmamaktadır.

Bilgilerinizi ve gereğini arz ederim.

Prof. Dr. Arif KARADEMİR
Rektör



T.C.
ÇANAKKALE ONSEKİZ MART ÜNİVERSİTESİ REKTÖRLÜĞÜ
Yabancı Diller Yüksekokulu Müdürlüğü



Sayı : E-12164519-302.08.01-2200075114
Konu : Anket İzni (Pınar ÇANKAYA)

04.04.2022

ÇANAKKALE ONSEKİZ MART ÜNİVERSİTESİ REKTÖRLÜĞÜNE
(Öğrenci İşleri Daire Başkanlığı)

İlgi : 31.03.2022 tarihli ve E-93130991-302.08.01-2200072690 sayılı yazınız.

Üniversitemiz Lisansüstü Eğitim Enstitüsü Yabancı Diller Eğitimi Anabilim Dalı İngiliz Dili Eğitimi Doktora Programı öğrencisi Pınar ÇANKAYA, "English Language Teachers' Assessment Literacy: Identifying Assessment Needs and Beliefs of EFL Teachers, Teacher Candidates and Teacher Educators" konulu tez çalışması kapsamında Yüksekokulumuzda ders vermekte olan öğretim görevlilerine anket/ölçek araştırmasına ilişkin anket izni uygun görülmüştür.

Bilgilerinizi ve gereğini arz ederim.

Prof. Dr. Aysun YAVUZ
Yabancı Diller Yüksekokulu Müdürü

Evrak Tarih ve Sayısı: 01.04.2022-190603



T.C.
ÇANAKKALE ONSEKİZ MART ÜNİVERSİTESİ REKTÖRLÜĞÜ
Öğrenci İşleri Daire Başkanlığı



Sayı : E-93130991-302.08.01-2200072658
Konu : Anket İzni (Pınar ÇANKAYA)

01.04.2022

DAĞITIM YERLERİNE

İlgi : Lisansüstü Eğitim Enstitüsü Müdürlüğünün 30.03.2022 tarihli ve E-95564340-302.08.01-2200071008 sayılı yazısı.

Üniversitemiz Lisansüstü Eğitim Enstitüsü Yabancı Diller Eğitimi Anabilim Dalı İngiliz Dili Eğitimi Doktora Programı öğrencisi Pınar ÇANKAYA, "English Language Teachers' Assessment Literacy: Identifying Assessment Needs and Beliefs of EFL Teachers, Teacher Candidates and Teacher Educators" konulu tez çalışması kapsamında Üniversitemiz Yabancı Diller Yüksekokullarında ders vermekte olan öğretim görevlilerine anket/ölçek araştırması yapmak istemektedir.

Gerekli iznin verilmesi hususunda gereğini arz ederim.

Prof. Dr. Pelin KANTEN
Rektör V.



T.C.
ÇANAKKALE ONSEKİZ MART ÜNİVERSİTESİ REKTÖRLÜĞÜ
Eğitim Fakültesi Dekanlığı



Sayı : E-68203582-605.01-2200079576
Konu : Anket Uygulama İsteminiz

12.04.2022

Sayın Pınar ÇANKAYA
Üniversitemiz Lisansüstü Eğitim Enstitüsü Doktora Öğrencisi
(pcankaya@nku.edu.tr)

“English Language Teachers’ Assessment Literacy: Identfying Assessment Needs and Beliefs of EFL Teachers, Teacher Candidates, and Teacher Educators” başlıklı araştırmanız kapsamında Fakültemiz İngilizce Öğretmenliği programı 4. sınıf öğrencilerine anket uygulama isteminiz Fakültemiz Bilimsel Araştırmaları Değerlendirme Kurulu tarafından incelenmiş ve uygun görülmüştür. Bilgilerinizi ve gereğini rica ederim.

Prof. Dr. Salih Zeki GENÇ
Eğitim Fakültesi Dekanı





T.C.
PAMUKKALE ÜNİVERSİTESİ
Öğrenci İşleri Daire Başkanlığı

Sayı : E-93282220-300-194021
Konu : Anket İzni (Pınar ÇANKAYA)

11.04.2022

ÇANAKKALE ONSEKİZ MART ÜNİVERSİTESİ REKTÖRLÜĞÜNE

İlgi : a) 01.04.2022 tarihli ve E-93130991-302.08.01-2200072658 sayılı yazınız.
b) Yabancı Diller Yüksekokulu Müdürlüğünüzün 09.02.2022 tarihli ve E-63788039-300-193515 sayılı yazısı.

Üniversiteniz Lisansüstü Eğitim Enstitüsü Yabancı Diller Eğitimi Anabilim Dalı İngiliz Dili Eğitimi Doktora Programı öğrencisi Pınar ÇANKAYA'nın "English Language Teachers' Assessment Literacy: Identifying Assessment Needs and Beliefs of EFL Teachers, Teacher Candidates and Teacher Educators" konulu tez çalışması kapsamında, Üniversitemiz Yabancı Diller Yüksekokulundaki İngilizce hazırlık sınıflarında ders veren öğretim elemanlarına İlgili (a)'da kayıtlı yazı Ek'inde gönderilen anket/ölçeği uygulama talebi Rektörlüğümüzce uygun görülmüştür.
Gereğini bilgilerinize arz ederim.

Prof. Dr. Ahmet KUTLUHAN
Rektör

Ek: İlgili (b) Yazı (1 Sayfa)

CURRICULUM VITAE

KİŞİSEL BİLGİLER

İsim SOYİSİM : PINAR ÇANKAYA
Doğum Yeri :
Doğum Tarihi :

EĞİTİM DURUMU

Lisans Öğrenimi :
Yüksek Lisans Öğrenimi :
Doktora Öğrenimi :
Bildiği Yabancı Diller :

BİLİMSEL FAALİYETLERİ

- a) Yayınlar
- b) Bildiriler
- c) Katıldığı Projeler

İŞ DENEYİMİ

Çalıştığı Kurumlar ve Yıl:

İLETİŞİM

E-posta Adresi :
ORCID :