



YEDITEPE UNIVERSITY

GRADUATE SCHOOL OF EDUCATIONAL SCIENCES

DESIGNING AN ONLINE LANGUAGE ASSESSMENT TRAINING COURSE
FOR IMPROVING LANGUAGE ASSESSMENT LITERACY OF SECONDARY
SCHOOL EFL IN-SERVICE TEACHERS IN MERSIN

A Thesis/Dissertation Submitted

by

Aslı Bağışlayıcı

In Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy
in
English Language Education

Supervisor

Dr. Zeynep Banu Koçoğlu

Istanbul - 2023

DESIGNING AN ONLINE LANGUAGE ASSESMENT TRAINING COURSE
FOR IMPROVING LANGUAGE ASSESSMENT LITERACY OF SECONDARY
SCHOOL EFL IN-SERVICE TEACHERS IN MERSIN

by
Aslı Bağışlayıcı

APPROVED BY:

Dr. Zeynep Banu KOÇOĞLU
(Yeditepe University)
(Thesis Supervisor)

Prof. Hossein FARHADY
(Yeditepe University)

Prof. Yasemin BAYYURT
(Boğaziçi University)

Dr. Evrim EVEYİK AYDIN
(Yeditepe University)

Dr. Merve SELÇUK
(Altınbaş University)

-

DATE OF APPROVAL:/....../2023

DECLARATION OF ORIGINALITY

I hereby declare that this thesis is my own work and that all information in this thesis has been obtained and presented following academic rules and ethical conduct. I have fully cited and referenced all material and results as required by these rules and conduct, and this thesis study does not contain any plagiarism. The necessary permissions have been obtained if any material used in the thesis requires copyright. No material from this thesis has been used to award another degree.

To the best of my knowledge and belief, it contains no material previously published or written by another person nor material accepted for the award of any other degree except where due acknowledgment has been made in the text.

I accept all kinds of legal liability that may arise in cases contrary to these situations.

Aslı Bağışlayıcı

Signature

ABSTRACT

DESIGNING AN ONLINE LANGUAGE ASSESMENT TRAINING COURSE FOR IMPROVING LANGUAGE ASSESSMENT LITERACY OF SECONDARY SCHOOL EFL IN-SERVICE TEACHERS IN MERSIN

Changes in education policies and movements in many countries have led to a high demand for a high level of assessment literacy. In order to prepare individuals to thrive in the 21st century, it is crucial to develop 21st century thinking skills, such as complex thinking, elaborated communication, collaboration, and problem-solving in authentic contexts. Quality education, instrumental for fostering these skills, relies on effective assessment. The purpose of this study is to assess language teachers' levels of assessment literacy knowledge and develop their skills through an in-service assessment workshop lasting six weeks. Prior to designing the workshop, the Professional Knowledge of Assessment Questionnaire (PKAQ), originally developed by Fulcher (2012); adopted by Farhady & Tavassoli (2018), was adapted to understand the language assessment needs of the participants. Two pilot studies were conducted. Following the results of the PKAQ, a focus group interview was conducted to gain a deeper understanding of the participants' assessment literacy needs. In addition to the results of the needs analysis and the findings of the interview, The researcher also examined the content pages of language assessment books used by instructors and professors working at education faculties and General Competence Indicators regarding Assessment suggested by MoNe were examined to develop a strong base for the workshop content. The workshop schedule was set. Participants took the Professional Knowledge of Assessment Test (PKAT) (Farhady, 2016) to understand their current level of assessment literacy before attending the workshop. PKAT was applied twice as pre- and post-tests of the study. After three weeks, halfway through the planned workshop duration, participants were invited to a face to face meeting to evaluate the workshop and gain deeper understanding about their actual experiences with the workshop. Participants were also sent three-question online form to evaluate the workshop individually so far. Considering the results of the meeting and the answers given to the online forms, necessary changes were made. Finally, the participants were given PKAT as post-test. The results of the study suggest that the participating

teachers had varying levels of understanding and knowledge regarding language assessment concepts, and that they had a clear need for professional development in this area. The workshop was found to have a positive impact on their language assessment literacy levels, with the teachers reporting increased confidence and competence in designing and administering assessments, as well as analyzing and interpreting assessment data.

Keywords: Assessment Literacy, Language Assessment Literacy



ÖZET

MERSİN'DE GÖREV YAPMAKTA OLAN İNGİLİZCE ÖĞRETMENLERİNİN ÖLÇME-DEĞERLENDİRME OKUR-YAZARLIKLARININ GELİŞTİRİLMESİ AMACI İLE ÇEVİRİMİÇİ DERS DİZAYN EDİLMESİ

Eğitim politikalarındaki değişiklikler ve hareketler bir çok ülkede yüksek seviye ölçme-değerlendirme okur-yazarlığı ihtiyacını doğurmaktadır. 21. yy. da bireylerin başarılı olması için 21. Yüzyıl düşünme becerilerinden kompleks düşünme, gelişmiş iletişim becerileri, işbirliği ve problem çözme yeteneklerini geliştirmeleri gerekmektedir. Sorunların üstesinden başarı ile gelebilen bireylerin yetişmesinde eğitimin rolü inkar edilemez ve bu da ancak yüksek kalite eğitim ile başarılabilir. Bu çalışmanın amacı, İngilizce öğretmenlerinin ölçme-değerlendirme okur-yazarlıklarının değerlendirilmesi ve bu konudaki becerilerinin ölçme-değerlendirme atölye çalışması yoluyla geliştirilmesidir. Söz konusu atölye çalışması dizayn edilmeden önce, Fulcher (2012) tarafından geliştirilen ve Farhady & Tavassoli (2018) tarafından uyarlanan Ölçme-Değerlendirme Mesleki Bilgi Anketi düzenlenmiştir. Ölçme-Değerlendirme Mesleki Bilgi Anketine ilişkin iki adet pilot çalışma düzenlenmiştir. Anketin uygulanmasının ardından, katılımcıların dilde ölçme-değerlendirme okur-yazarlığına ilişkin ihtiyaçlarının daha detaylı belirlenmesi amacıyla gönüllü katılımcılar ile çevrimiçi odak grup görüşmesi yapılmıştır. Araştırmacı, çevrimiçi atölye çalışmasına ilişkin programı hazırlamadan önce ayrıca Eğitim Fakülteleri İngilizce Öğretmenliği Bölümlerinde kullanılmakta olan Ölçme-Değerlendirme ders kitaplarının içeriklerini ve Milli Eğitim Bakanlığı tarafından geliştirilen Öğretmenlik Mesleği Genel Alan Yeterlikleri kapsamında yer alan Ölçme-Değerlendirme Yeterlik göstergelerini de incelemiştir. Atölyeye ilişkin ders programı oluşturulmuştur. 6 hafta süren atölye çalışması başlamadan önce, Ölçme-değerlendirme Mesleki Bilgi Testi (Farhady, 2016) ön test olarak katılımcılara uygulanmıştır. Üç hafta sonunda, atölye çalışmasının değerlendirilmesi amacıyla katılımcılar yüzyüze bir toplantıya davet edilmişlerdir. Toplantıya ek olarak, katılımcılara üç soruluk değerlendirme formu gönderilmiştir. Gerek yüzyüze gerçekleştirilen toplantı, gerekse katılımcıların bireysel olarak doldurmuş oldukları formlar dikkate alınarak düzenlemeler yapılan atölye çalışmasının ardından katılımcılara Ölçme-Değerlendirme Mesleki Bilgi

Testi son test olarak uygulanmıştır. Çalışma sonucunda, katılım sağlayan öğretmenlerin çeşitli seviyelerde ölçme-değerlendirme okuryazarlığına sahip oldukları ve mesleki gelişim ihtiyaçları olduğu görülmektedir. Çalışma kapsamında gerçekleştirilen çalıştayın yabancı dil ölçme değerlendirme okur-yazarlıkları üzerinde pozitif etki gösterdiği tespit edilmiştir. Ayrıca katılım sağlayan öğretmenler çalışma sayesinde özgüvenlerinde ve ölçme değerlendirme bileşenleri (ölçme-değerlendirme araçlarının dizayn edilmesi ve uygulanması; ölçme-değerlendirme sonuçlarının analiz edilmesi ve sonuçların değerlendirilmesi) konusunda ilerleme kaydettikleri bulgular arasında yer almaktadır. Çalışma sonucunda, katılımcı öğretmenlerin, ölçme-değerlendirme konusundaki yeterliklerinin çeşitli seviyelerde olduğu ortaya çıkmış olup; mesleki olarak bu alanda gelişime ihtiyaçları olduğu ortaya konmuştur. Çalışma kapsamında düzenlemiş olan çevrimiçi atölye çalışmasının katılımcıların ölçme-değerlendirme okuryazarlık seviyeleri üzerinde olumlu etki gösterdiği anlaşılmıştır. Bunlara ek olarak, öğretmenlerin özgüven ve yeterliklerinin test dizayn etme ve ölçme-değerlendirme süreçlerini yönetebilmelerinin yanı sıra değerlendirme verilerinin analiz edilmesi ve değerlendirilmesi konusunda yeterlik ve güven geliştirdikleri gözlemlenmiştir.

Anahtar Kelimeler: Ölçme-Değerlendirme Okur Yazarlığı, Dilde Ölçme- Değerlendirme Okur Yazarlığı

DEDICATION

To children, angels of the earth...



ACKNOWLEDGEMENTS

This was a long journey for me that had started with hope. Now, it is time to enlighten the way for others who will begin the same journey.

I would like to express my sincere gratitude to those who have helped me in completing this dissertation. I would first like to thank my supervisor, Dr. Zeynep Koçoğlu for her professional guidance, ongoing support, helpful criticism and sincere attitude throughout the development of this research. She has inspired me to do research on language assessment literacy and has actively engaged me in this process by providing me with the support to be both the researcher despite working in an another field. I am particularly indebted to her for being more than a supervisor and for being such an approachable person, without which writing a dissertation would turn into an endless journey. My deepest feelings of gratitude also go to my second supervisor, Prof. Hossein Farhady for his encouragement and trust in me. He has been by my side for ten years, and I am really grateful to be his graduate student. My sincere and special thanks go to Prof. Yasemin Bayyurt for always being so responsive to my need for help. Her invaluable feedback and comments helped me continue my study even in very challenging situation. My appreciation is also to Dr. Evrim Eveyik Aydın for patiently and calmly responding to my never-ending questions.. Without the hours we spent talking, on the phone, this study would not be completed yet. I feel very lucky to have her in my life. My branch manager, Eylem Eren, also deserves special thanks as she has been always there for me in any difficulty I face in my life and during my studies. I should also acknowledge the contributions of Feyza Tüzün, Tuğçe Arkan, Süleyman Burak Topaktaş, Göktürk Şener and Aysun Aldı who unconditionally supported me and always encouraged me during my academic studies. I am very proud to have them in my life. Special thanks also go to Mustafa Berat Yerman, head of the IT department, Mersin Metropolitan Municipality, without whom support the study would not finish. I would also like to express my appreciation to my six volunteer teachers for their willingness to participate in the study. Their constant feedback, comments and support were the basis of designing and continuing this program. Without their agreement on taking part in the study, this thesis would not be complete. My wholeheartedly thanks go to a great team, from Directorate of Provincial National Education, Center of Measurement and Assessment. Their support and

collaboration cannot be defined through words. They are the gift of the life that hold my motivation throughout the study.

My thanks from the deepest corner of my heart go to my family. My mother and father, Filiz and Süleyman Bağışlayıcı, for their never ending support, sacrifices, trust and unconditional love without which I would have lost hope. My brother Ali deserves special thank from me. I am grateful his always being there when I needed to cheer up. My husband, Engin Habib Özcan, for creating a peaceful and supporting atmosphere at home, for being a great husband and an ideal father, without whose support, I would not have succeeded in this journey. And finally, my four year old son, Ali Özcan, who is the light of my way even in the darkest night. Life has gained another dimension with his birth. I appreciate his existence every single moment.

Last but not the least, my uncle and my aunt, Murat and Ayla Adak who behaved me as their daughter during my stay at their cozy and motivating home, without them I would not dare to start my Phd studies. Countless thanks to them, they will also have a special place in my life.

Thank you all...

TABLE OF CONTENTS

DECLARATION OF ORIGINALITY	iii
ABSTRACT.....	iv
ÖZET	vi
DEDICATION	viii
ACKNOWLEDGEMENTS	ix
TABLE OF CONTENTS	xi
LIST OF FIGURES	xiv
LIST OF TABLES	xv
LIST OF ABBREVIATIONS	xvi
1. INTRODUCTION	17
1.1. Background of the Study.....	17
1.2. The Theoretical Basis for the Study	18
1.3. Statement of the Problem.....	20
1.4. Significance of the Study	22
1.5. The Role of the Researcher	22
1.6. Research Questions.....	23
1.7. Outline of the Study	23
1.8. Definition of Terms	23
2. LITERATURE REVIEW	25
2.1. Literacy and Assessment	25
2.2. Language Assessment Literacy.....	27
2.3. Teachers' Language Assessment Literacy.....	31
2.3.1. Teachers' Language Assessment Literacy and Training	32
2.3.2. General Competencies of Teaching Profession in Turkey	37
2.4. Teachers' Language Assessment Literacy Perceptions.....	41
2.5. Empirical Studies on Language Assessment Literacy	44
2.6. Chapter Summary	46

3. METHODOLOGY	49
3.1. Setting	49
3.2. Participants.....	50
3.2.1. Participants of the Pilot Study for PKAQ	50
3.2.2. Participants of PKAQ	51
3.2.3. Participants of the Online Workshop, Interviews and PKAT	51
3.3. Research Design	52
3.4. Instruments.....	53
3.4.1. Professional Knowledge of Assessment Questionnaire (PKAQ)	54
3.4.2. Professional Knowledge of Assessment Test (PKAT)	54
3.4.3. Semi-structured Interviews	56
3.4.4. Reflective Notes.....	57
3.4.5. General Competency Indicators for English Language Teachers	58
3.5. Data Collection Procedure	58
3.5.1. Language Assessment Literacy Workshop	59
3.6. Data Analysis.....	64
3.7. Chapter Summary	67
4. RESULTS	68
4.1. Findings of the Professional Knowledge of Assessment Questionnaire.....	68
4.2. Findings on Language Assessment Needs.....	74
4.3. Findings on the Effect of Workshop on LA levels	78
4.4. Findings on the Reflective Notes.....	82
4.5. Findings on the Reflective Notes of Basics of Testing.....	82
4.6. Findings on the Reflective Notes of Functions of Assessment	83
4.7. Findings of the Reflective Notes on Psychometrics Features of a Test	85
4.8. Findings on the Reflective Notes on Assessing Language Skills.....	92
4.9. Findings on The Reflective Notes of Assessing Reading Skills	92
4.10. Findings on The Reflective Notes of Assessing Speaking Skills	94
4.11. Findings on The Reflective Notes of Assessing Writing Skills.....	96
4.12. Findings on The Reflective Notes of Assessing Listening Skills	98
4.13. Findings on The Reflective Evaluatuation of the Workshop.....	100
4.14. Effect of Online Workshop on Development of Competency Indicators ...	101

4.15. Chapter Summary	103
5. DISCUSSION AND CONCLUSION	104
5.1. Analysis of Workshop Efficacy.....	105
5.2. In-Depth Exploration of Diversity in Learning Progression	106
5.3. Comprehensive View on Assessment Literacy Needs.....	107
5.4. Detailed Exploration of Workshop Structure and Content.....	107
5.5. Pedagogical Implications.....	109
5.6. Limitations.....	112
5.7. Recommendations for Further Research	113
REFERENCES:	115
APPENDIX A: INFORMED CONSENT FROM.....	127
APPENDIX B: ETHIC COMMITTEE REPORT	128
APPENDIX C: PROFESSIONAL KNOWLEDGE OF ASSESSMENT QUESTIONNAIRE	129
APPENDIX D: PROFESSIONAL KNOWLEDGE OF ASSESSMENT (PKA) TEST	132
APPENDIX E: PKAQ ITEM ALIGNMENT TABLE	140
APPENDIX F: THEORETICAL FRAMEWORK OF LAL COMPONENTS	142
APPENDIX G: SUGGESTIONS.....	143
APPENDIX H: PROMPTS FOR REFLECTIVE NOTES.....	144
APPENDIX I: WORKSHOP SYLLABUS.....	146
APPENDIX J: PKAQ CFAs with 4 FACTOR SOLUTION.....	152

LIST OF FIGURES

Figure 1 Levels of AL/LAL differentiated according to stakeholders' constituency (Taylor, 2013, p. 409)	29
Figure 2 Differentiated AL/LAL profiles for four constituencies	30
Figure 3 Eight components of LAL for language teachers	36
Figure 4 General competencies for teaching profession (MoNE, 2017)	40
Figure 5 General competencies for assessment and evaluation (MoNE, 2017)	41
Figure 6 General competencies of teaching profession: Assessment and evaluation (MoNE, 2017)	58
Figure 7 Procedure for developing online language assessment literacy workshop	60
Figure 8 Psychometric features of a test (Çeiik, Çelik and Coombe, 2021, p. 44).....	86
Figure 9 Psychometric features of test (Çelik, Çelik and Coombe, 2021, p. 45).....	87
Figure 10 Psychometric features of a test (Çelik, Çelik and Coombe, 2021, p.51).....	89
Figure 11 Psychometric features of a test (Çelik, Çelik and Coombe, 2021, p. 51).....	90
Figure 12 Psychometric features of a test (Çelik, Çelik abd Coombe, 2021, p. 47).....	91

LIST OF TABLES

Table 1 Theoretical framework for language assessment literacy.....	36
Table 2 Distribution of PKAT items and language assessment literacy topic.....	55
Table 3 Alignment between the PKAT items and the theoretical framework for LAL	56
Table 4 Summary of research design.....	66
Table 5 Cronbach alpha values for PKAQ	69
Table 6 Percentage distribution for PKAQ items 10, 11 and 12	72
Table 7 Comparison of pre- and post-test results for LAL course	79

LIST OF ABBREVIATIONS

AL	Assessment literacy
CoHE	Council of Higher Education
ELT	English Language Teaching
LAL	Language assessment literacy
MoNE	Ministry of National Education
PKAQ	Professional Knowledge of Assessment Questionnaire
PKAT	Professional Knowledge of Assessment Test

1. INTRODUCTION

This chapter presents the background of the study, the theoretical basis for the study, statement of the problem, significance of the study, the role of researcher, research questions, outline of the study and definition of terms.

1.1. Background of the Study

Since the establishment of the Turkish Republic, teacher education has been a crucial issue. The dynamic nature of teaching and societal needs have necessitated many changes and reforms throughout history. Teachers perform their profession in a constantly changing social and cultural environment. These changes have a significant impact on education, and as a result, educational programs need to be regularly updated and reformed. Recent studies have identified the following expectations from contemporary teacher education programs: taking responsibility for school and student needs, integrating theoretical and pedagogical knowledge, being information literate, having research competence, being motivated for self and professional development, collaborating with colleagues, parents, and society, creating a conducive learning environment that promotes higher order thinking skills, and adapting teaching methods to meet the diverse needs of students from different backgrounds (MoNE, 2017).

After the role of teacher education being assigned to the universities, a comprehensive study on Teacher Education Programs was conducted under the "Project for Improving National Education" between 1996 and 1997. In response to the announcement of mandatory 8-year education, the Council of Higher Education (YÖK) redesigned the Curriculum of Education and Educational Sciences Faculties. The first assessment-related courses were integrated into the ELT Curriculum in 1998, with two more programs introduced in 2007 and 2009, respectively. These changes were made in response to the dynamic nature of teaching and societal needs, aiming to equip teachers with the necessary skills and knowledge to meet the contemporary expectations of the teaching profession.

The curriculum of educational sciences and faculties of education have undergone significant changes due to the contemporary developments in the field of educational sciences, structural changes in the Turkish Education System, and societal needs and demands. Therefore, the update of teacher education programs has become mandatory. In accordance with this, the General Board of the Council of Higher Education made a decision on 28.02.2017 to revise the template for educational sciences and faculties.

The Ministry of National Education's (MoNE) introduction of a new ELT curriculum highlighted the need for changes in teacher education programs. Additionally, the Bologna Process brought up issues related to adaptation, quality, and accreditation. The Council of Higher Education conducted accreditation and quality studies, while updates were made to the General Professional Competency Areas in 2017.

The recent change in the teacher education programs also resulted in the transformation of the Testing and Assessment courses in English language Teaching curriculum. English Teacher Education programs include a course specifically designed for developing and evaluating tests for English Education. The course covers different tests and measurement methods for various age groups and levels of language competency, basic principles for measuring and evaluating language skills such as reading, writing, listening, speaking, vocabulary knowledge, and grammar level. It also provides techniques for test preparation, evaluation criteria, and guidelines for preparing test items and test evaluation.

Despite recent restructuring in the Education Faculties and curriculum of ELT Teacher Education Programs, the need for sustained professional development in language assessment remains critical. The dynamic nature of teaching demands that teachers keep up with the changes in education, and in-service training is an indispensable part of this development. Teacher education should be seen as an ongoing, dynamic system (CoHE, 2007).

1.2. The Theoretical Basis for the Study

The word pedagogy that comes from *peda* (*paid* meaning *child* in Greek) and *agagous* (meaning *leader of*) apparently support the idea of teachers' being key agent in key social

processes. Throughout the process of learning, each individual is developing a social identity to fulfil a significant public role (Deulen, 2013). Unlike the major limitations of humanistic theory and constructivism, which solely focus on individuals, social constructivist theory put emphasis on the vitality of social context in learning (Roberts, 1998). Shift in learning theory to social constructivist theory promotes the significance of Assessment for learning as an instructional intervention. This influences the knowledge base that teachers need to develop in order to make appropriate instructional decisions that will lead learner to fulfil a social role in their future careers.

Since teachers are the leaders of the change in individuals' life, their role of making education decisions become essential for individuals to develop appropriate position in social life. Aiming to develop language assessment literacy levels of language teachers in order them to be more competent in making appropriate educational decision to assist their students' future career, this study is theoretically based on social constructivist learning theory. According to Social Constructivist Learning Theory, learning takes place in a social context. Vygotsky (1978) introduced the construct of the zone of proximal development (ZPD), which is what children can do with assistance. It is the distance between the actual development level and the development level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers. In summary, productive interactions are those that orient instruction toward the ZPD (cited in Palinscar, 1998). It is apparent that one of the basic construct of Social Constructivist Theory, Zone of Proximal Development constitutes the base for assessment for learning which can be mentioned as the reflection of social constructivist theory into the language assessment practices for this study. According to Lian et al. (2014), Assessment for learning does not mean performing assessment practices with the aim of giving grades to the students. On the contrary to this, as an intervention, Assessment for Learning requires teachers to use assessment practices constantly in order to contribute students' learning process (DeLuca & Volante, 2016).

Since the ultimate aim of this study is to enhance language assessment literacy skills of ELT teachers through which learning processes of the students will be improved that will contribute to the future careers of the learner, theoretical framework of this study is based on Social Constructivist Theory.

1.3. Statement of the Problem

Assessment is an essential component of the teaching cycle, and its results can have a long-lasting impact on learners' academic lives. Therefore, possessing adequate knowledge of assessment and a high level of assessment literacy play a fundamental role in the professional development of teachers. However, according to Gelbal and Kelecioğlu (2007) and Çakan (2004), both in-service and pre-service teachers often lack the ability to use assessment methods effectively. Despite the crucial role of assessment in students' educational journey, many teachers feel inadequately prepared for assessment (Koh, 2011; Koh et al., 2017; Lam, 2015; Popham, 2011).

Despite the curriculum regulations for teacher education programs, many preservice teachers express concerns about the insufficiency of assessment course content and lack of opportunities to practice theoretical knowledge of assessment (Büyükkarcı, 2016; Hatipoğlu, 2015; Jin, 2010; Jeong, 2013; Lam, 2015; Mede & Atay, 2017; Öz & Atay, 2017; Şahin, 2015; Şahin, 2019). Consequently, it is suggested that most language teachers gain experience in using theoretical knowledge of assessment as they begin their teaching careers (Arias & Maturana, 2005; Cheng & Rogers & Hu, 2004; Diaz, Alarcon & Ortiz, 2012; Frodden, Restrepo & Maturana, 2004; Lopez & Bernal, 2009; Munoz, Palacio & Escobar, 2012; Volante & Fazio, 2007; Yastıbaş, 2018).

Studies have shown that language teachers with limited knowledge of assessment tend to rely more on summative assessment methods (Lopez & Bernal, 2009). This lack of competency can result in language teachers struggling to manage the results of standardized tests (Ghafar et al., 2013; Kimors, Abdolmehdi, & Naser, 2011; Leap et al., 2015; Xu & Brown, 2017). On the other hand, language teachers who possess sufficient language assessment literacy are successful in applying various assessment methods effectively (Herrera & Marcias, 2015).

Studies have shown that many teachers are interested in using alternative and modern assessment tools, but they are often limited by heavy workloads and time constraints (Gönen & Akbarov, 2015; Munoz et al., 2012; Rogers et al., 2007; Shohamy et al., 2008). Despite their interest, many teachers still tend to use traditional assessment methods due to these constraints,

as well as external factors such as government policies on language teaching curriculum, course content, and objectives that may not emphasize the use of alternative methods for assessing learning (Jannatti, 2015).

Teachers' beliefs and professional knowledge base also influence their choice of assessment methods. However, preservice language teacher education often lacks sufficient training in assessment and testing, leading language teachers to express their assessment-related needs in various studies (Tsagari & Vogt, 2014). Even though many language teachers are competent in language assessment, they still require additional training to improve their knowledge base (Kavaklı & Arslan, 2019; Ölmezer-Öztürk & Aydın, 2019; Sarıyıldız, 2018; Semiz & Odabaşı, 2016; Vogt & Tsagari, 2017). Consequently, short-term assessment courses have been suggested as a means to bridge the gap between assessment and instruction, promoting language teachers' professional development (Mahapatra, 2016; Nier et al., 2013; Riestenberg et al., 2010).

In today's world, educational policies and movements in many countries prioritize assessment literacy to develop individuals capable of thriving in the 21st century. To succeed in this era, individuals require 21st century thinking skills, such as complex thinking, effective communication, collaboration, and problem-solving in real-life situations. High-quality education is essential to achieve this goal and its impact on the development of future individuals cannot be denied. Therefore, language teachers' assessment literacy knowledge plays a significant role in providing quality education to learners. An online language assessment literacy workshop for in-service EFL teachers should be designed to enhance their knowledge and skills in this regard. While assessment is an integral part of teaching and learning, research in this area is still in its infancy. To the best of our knowledge, no study has been conducted to understand the language assessment needs of in-service language teachers and then design an online workshop to enhance their assessment literacy levels. Hence, this study aims to fill this important gap by investigating language teachers' levels of assessment literacy knowledge and develop their skills through an in-service assessment workshop lasting six weeks.

1.4. Significance of the Study

Despite the existence of research on the development of assessment literacy in core subjects such as Maths, Social Sciences, Science, and Turkish (e.g. Cambay, 2016; Gul, 2011; Karaman, 2014; Yalvac, 2016), little attention has been paid to the assessment literacy levels of language teachers in Turkey. Based on the Higher Education Institution's database, only four studies focus on English language preservice teachers and six studies focus on EFL instructors. Moreover, only two studies investigate the assessment literacy levels of in-service EFL teachers. One of these studies is a MA thesis by Ballıdağ (2020) on “Exploring the Language Assessment Literacy of Turkish in-service EFL teachers” and a Phd dissertation by Sarıyıldız-Canlı (2023) on “Language Assessment Literacy Level and Needs of In-Service EFL Teachers in Turkey.”

This study is unique compared to all previous research on this topic in terms of target population and the method used to develop the language assessment literacy levels of in-service language teachers. Additionally, no similar study has designed an online workshop specifically aimed at addressing the language assessment literacy needs of in-service language teachers. Therefore, this study provides empirical evidence for university researchers, district office personnel, school administrators, and classroom teachers.

1.5. The Role of the Researcher

In this study researcher is the key actor in defining the research questions, collecting, processing and analyzing data. Also, the researcher has a major role in interpreting the results of the data analysis and ensuring the alignment of the all data collection instruments with the theoretical framework of the study and the general competencies of teaching profession suggested by MoNE.

In addition these, researcher is also the moderator of the online workshops designed for improving Language Assessment Literacy competencies of the participating in service teachers. Throughout this process, the syllabus for workshop, tasks and course evaluation steps were also carried out by the researcher.

1.6. Research Questions

The research questions of the study are the following:

- 1- What are the language assessment literacy needs of EFL in-service teachers?
- 2- What are the competency indicators that EFL in-service teachers have before the language assessment literacy (LAL) workshop?
- 3- What competency indicators do EFL in-service teachers acquire through LAL workshop?
- 4- Do LAL workshop effectively improve the language assessment literacy competencies of EFL in-service teachers?

1.7. Outline of the Study

This dissertation is structured into five chapters, beginning with an introduction that sets the stage with background information, problem statement, study significance, and an overview of the methodology and key terms. The second chapter delves into a comprehensive literature review, focusing on aspects such as literacy and assessment, Language Assessment Literacy, and empirical studies in this field. The methodology, detailed in the third chapter, describes the research design, theoretical framework, participants, and data collection instruments. Chapter four presents the study's results, analyzing the data gathered. Finally, the fifth chapter ties together the conclusions and implications of the study, discussing its limitations and suggesting avenues for future research, ending the dissertation with references, appendices, and figures.

1.8. Definition of Terms

Assessment: Assessment can be defined as “the process of collecting information about something and the outcome of this procedure” (Bachman & Palmer, 2010, p. 20).

Language Assessment: The process of “collecting information about language learner’s ability through systematicity and substantive grounding” (Bachman & Palmer, 2010, p.20).

Evaluation: Evaluation is described as “the activity that involves making judgements about people or events considering basic information gathered through tests and assessments” (Bachman & Palmer, 2010, p.21 ; Douglas, 2010, p. 5).

Testing: “Process of measuring a person’s ability knowledge or performance in a given domain” (Brown, 2004, p. 3).

Assessment Literacy: Assessment literacy largely has been defined as “teachers’ understanding of assessment processes as well as their capacities to design assessment tasks, develop adequate criteria for making valid judgments on the quality of students’ performances, and understand and act upon the information that is collected through assessment” (Hay & Penney, 2013, p. 69–70).

Language Assessment Literacy: LAL incorporates knowledge of language, principles, and skills of language testing (Davies, 2008; Fulcher, 2012; Inbar-Lourie, 2013b).

2. LITERATURE REVIEW

This chapter provides an overview of the theoretical foundations of this dissertation and the related literature on language assessment studies, including studies conducted in Turkey and around the world. The chapter begins with a discussion of the rise of the term assessment literacy and continues with the process of defining the term Assessment Literacy in the field of Language Teaching.

2.1. Literacy and Assessment

Literacy is commonly defined as the ability to read and write, but it can also refer to competence and knowledge in a specific area (Koh, 2011). Due to the increasing responsibilities of teachers in assessment, assessment literacy has become increasingly important. However, teachers' assessment literacy should not be mistaken for mere knowledge of tests and the interpretation of results. To be assessment literate, one needs to be proficient in designing and using high-quality classroom assessments.

Assessment literacy was first coined by American Federation of Teachers in 1990. Although the term assessment literacy was not used directly, the competency areas had been defined as “selecting assessments, developing assessments for the classroom, administering and scoring tests, using scores to aid instructional decisions, communicating results to stakeholders, being aware of inappropriate and unethical uses of tests” (cited in Fulcher, 2012, p.115-116)

Building on this definition, Stiggins (1991, 1997, cited in Fulcher, 2012) provided a comprehensive definition of assessment literacy as "the range of necessary assessment knowledge and skills required for the new era of education." Stiggins further elaborated on this definition by emphasizing what an assessment literate stakeholder should know. According to him, an assessment literate individual should have a clear understanding of what they are assessing, why they are conducting the assessment, and what other types of assessments are available and appropriate.

In line with the definition of Stiggins summarized above, Hoyt (2005) and Paterno (2001) defines assessment literacy as teachers' competence in both basic principles of assessment and alternative assessment methods. Paterno (2001) also puts emphasis on the significance of quality of assessment. Siegel and Wissehr (2016) stress the teachers' competence in practicing assessment tools besides their theoretical knowledge of assessment.

Researchers have provided various definitions of assessment literacy. Upon examining these definitions, it becomes clear that certain common points are mentioned in each definition. These include a focus on classroom assessment, the importance of improving teachers' own teaching to enhance student learning, and the need for both theoretical and practical knowledge of assessment. Assessment literacy largely has been defined as “teachers’ understanding of assessment processes as well as their capacities to design assessment tasks, develop adequate criteria for making valid judgments on the quality of students’ performances, and understand and act upon the information that is collected through assessment” (Hay & Penney, 2013, p. 69–70). Therefore, assessment literacy is not just about having knowledge of assessment, but also about possessing the skills and competencies related to testing and classroom assessment. Although there is no universally accepted definition of assessment literacy, this study will define it as teachers' competence in selecting or creating appropriate assessment tools, using these tools to provide feedback on both students' learning and their own teaching, and thus improving their instructional practice.

It is widely acknowledged that having a high level of assessment literacy is crucial for providing quality education to learners. As such, this issue is of great concern to a wide range of stakeholders, including policy makers and pre-service teachers. However, despite the importance of assessment literacy, there is still no consensus on the specific components that make up this term. Even with the available resources and principles on assessment, satisfactory identification of these components has yet to be achieved (Brindley, 2001; Brown & Bailey, 2008; Davies, 2008 cited in Fulcher, 2012).

2.2. Language Assessment Literacy

Just like Assessment Literacy, Language Assessment Literacy (LAL) is a concept that lacks a clear definition and requires further research. LAL encompasses the knowledge, principles, and skills necessary for language testing (Davies, 2008; Fulcher, 2012; Malone, 2008). It has been suggested that many of the definitions developed thus far are based on the framework proposed by Brindley (2001 cited in Fulcher, 2012; Giraldo, 2018; Giraldo, 2021), which consists of five areas that language teachers should receive training in “the social context of assessment, defining and describing proficiency, constructing and evaluating language tests, assessment in the curriculum, putting assessment into practice”.

The debate over the definition of language assessment literacy was initiated by Davies (2008) who argued that skills, knowledge, and principles are the three essential elements of the assessment literacy construct. However, some argue that this construct is mainly concerned with college instructors and has nothing to do with K-12 teachers. On the other hand, LAL incorporates knowledge of language, principles, and skills of language testing (Davies, 2008; Fulcher, 2012; Inbar-Lourie, 2013b).

Fulcher (2012) is one of the most significant researchers who explains the rise of language assessment literacy. He suggests that the term language assessment literacy started to gain importance during the first decade of the 21st century due to the following three primary reasons:

Accountability: Fulcher explains that policy makers often use test results to achieve political goals, as tests are a quicker and easier solution than changing the curriculum for any reason. Additionally, externally mandated test results can be used to "raise educational standards," making accountability of tests crucial.

International Mobility: With language tests now being part of immigration policy, it is increasingly common for teachers to offer test preparation classes aimed at helping their students achieve high scores on large-scale language tests that are used to determine international mobility.

Assessment for Learning: As the role of assessment for learning gains importance in classroom practices for enhancing learning, it is expected that language teachers will need to develop competence in various assessment strategies to effectively apply classroom assessment and evaluate its success.

According to him, Language Assessment Literacy includes:

The knowledge, skills and abilities required to design, develop, maintain or evaluate, large-scale standardized and/or classroom based tests, familiarity with test processes, and awareness of principles and concepts that guide and underpin practice, including ethics and codes of practice. The ability to place knowledge, skills, processes, principles and concepts within wider historical, social, political and philosophical frameworks in order to understand why practices have arisen as they have, and to evaluate the role and impact of testing on society, institutions, and individuals. (Fulcher, 2012, p.125)

Koh et al. (2017) emphasizes that language teachers have increased responsibilities in testing and assessment, and having a high level of competence in both interpreting large-scale tests and classroom assessment is crucial. They also emphasize that the success of curriculum reforms and assessment-related policies relies on language assessment literacy. This sentiment is echoed by Luke (2011), as cited in Koh et al. (2017).

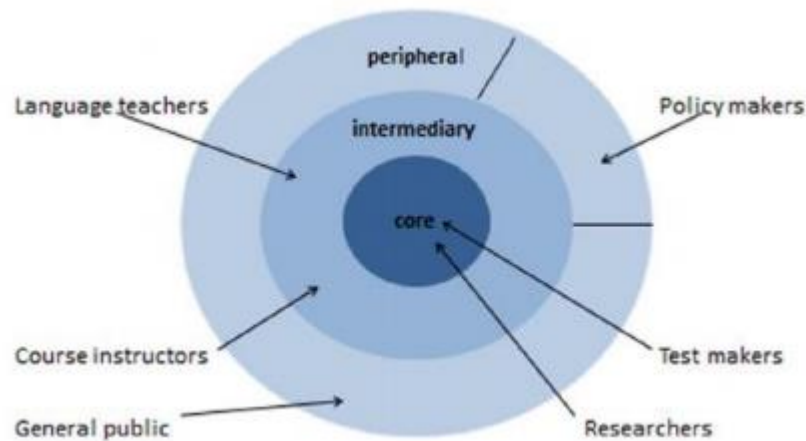
Davies (2008) describes Language Assessment Literacy as including a range of competencies, comprising skills, knowledge, and principles. Skills include areas such as test design, understanding test statistics, developing software for test delivery, and reporting. Knowledge relates to an understanding of language-related backgrounds, while principles focus on ethics and validity in language assessment. In contrast, Inbar-Loria (2008, 2013) offers a different viewpoint, suggesting that skills and knowledge form the foundational layer of a pyramid structure. At the apex of this pyramid are language-related components, including language acquisition, learning theories, and methods and approaches to language teaching and learning. Additionally, the top of the pyramid also covers the knowledge and skills necessary for teaching language skills to students of various ages and proficiency levels.

While Brindley (1997), Davies (2008), Fulcher (2012), O'Loughlin (2013), Vogt and Tsagari (2014) focus on the knowledge and skills that Language Assessment Literate stakeholders should have, Taylor (2013) proposes a model that stresses differential levels of language assessment-related expertise due to differing needs of different stakeholders. Taylor suggests that the degree of expertise differs among stakeholders.

In his model shown in Figure 1 "The circles expand outwards from an 'expert core' of assessment knowledge, skill and principles, with each successive ring [.....] representing the level of content/input that is required to meet the needs of a particular set of stakeholders" (Taylor, 2013, pp. 409).

Figure 1

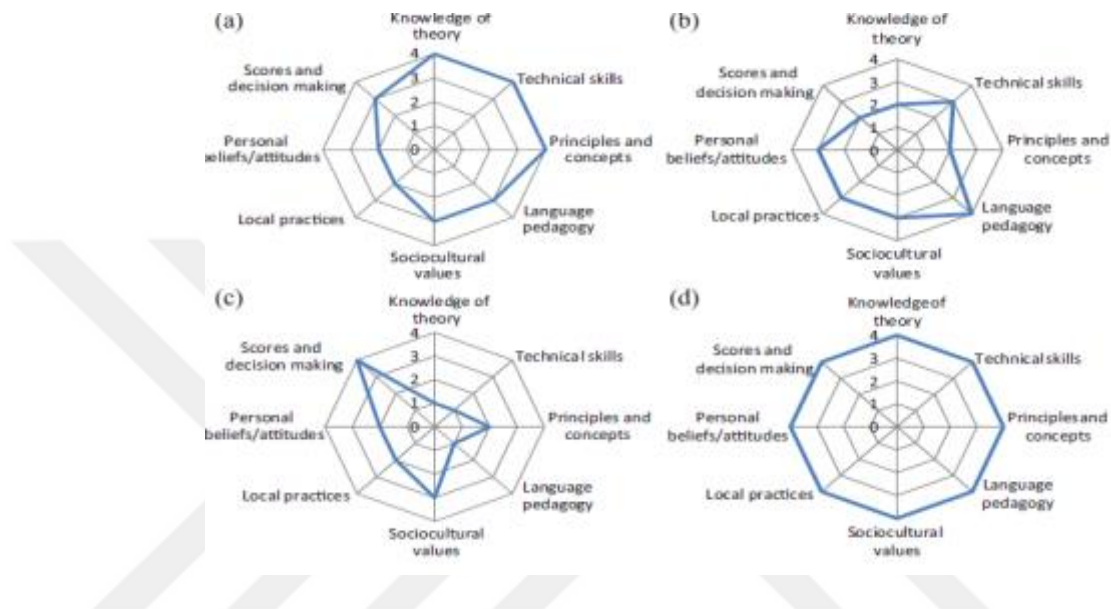
Levels of AL/LAL differentiated according to stakeholders' constituency



In order to give a detailed portrait of differing competency and knowledge levels of different stakeholders, Taylor proposed a model (Figure 1) in which eight skills and knowledge areas (Knowledge of theory, technical skills, principles and concepts, language pedagogy, sociocultural values, local practices, personal beliefs, scoring and decision making) are matched with numerical values (0-4) corresponding 5 proficiency levels suggested by Pill and Harding (2013) as shown in Figure 2.

Figure 2

Differentiated AL/LAL profiles for four constituencies



In this figure, “profile of test writer, profile of classroom teachers, profile of university administrators and profile of professional language testers” (Taylor, 2013, p. 401) are emphasized that the level of assessment literacy vary according to the differentiated profiles. In the assessment environment, each stake holder is expected to have differing level of competencies regarding the components of assessment literacy defined. For instance, the diagram b suggests that classroom teacher should have knowledge level 2 regarding principles and concepts, however the knowledge levels of professional language testers should be at 4 for all components of language assessment literacy. In their 2013 study, Phil and Hardig proposed a framework categorizing literacy into five distinct proficiency levels. At the lowest level, “Illiteracy,” individuals show a complete ignorance of language assessment concepts and methods. The next level, “Nominal Literacy,” is characterized by a basic understanding that certain terms are related to assessment, though this often comes with misconceptions. ‘Functional Literacy’ represents a sound grasp of basic terms and concepts in the field. The fourth level, “Procedural/Conceptual Literacy,” goes further, indicating not only an understanding of central concepts but also the application of this knowledge in practical scenarios. Finally, the highest level, “Multidimensional Literacy,” encompasses a comprehensive knowledge that

extends beyond ordinary concepts to include philosophical, historical, and social dimensions of assessment.

Despite Taylor's valuable contribution to the field of Language Teaching Assessment by specifying the differing needs of various stakeholders, two arguments have been raised. First, Harding and Kremmel (2016) argue that proficiency levels may vary from culture to culture and institution to institution, and therefore, categorizing them as black and white may not reflect the dynamic nature of stakeholders' needs. Instead, proficiency levels should be detailed by considering skills and knowledge corresponding to each level of proficiency. Secondly, proficiency levels corresponding to each competency level are not clear-cut, and may even vary among language teachers within the same institution.

The present study therefore defines language assessment literacy as the combination of language assessment-related knowledge and skills that pertain to the purpose of assessment and the utilization of assessment outcomes to improve instruction, resulting in a positive impact on learners' social and professional lives. In simpler terms, it refers to language teachers' familiarity with fundamental testing concepts, their ability to apply them in classroom instruction, and their proficiency in language assessment-related tasks (Inbar-Lourie, 2008; Melone, 2013; Taylor, 2009).

2.3. Teachers' Language Assessment Literacy

Since the term "Language assessment literacy" was first coined, it has caught the attention of the field and many studies have been carried out to define its components, current place among various stakeholders, language teachers' levels, beliefs, and perceptions of LAL, as well as the reasons behind low LAL levels and LAL-related training needs (Feretti et al. (2021), Firoozi et al. 2019; Fulcher, 2012; Levi & Inbar-Lourie, 2020; Sultana, 2019; Tsagari & Vogt, 2017; Vogt et al., 2020; Vogt & Tsagari, 2014; Xu & Brown, 2017).

Despite the evolving understanding of its significance on teaching practices and learning procedures, Language assessment literacy is still in its infancy. Various studies have

investigated the reasons for the low language assessment literacy levels of language teachers. Lack of assessment theory knowledge, low levels of assessment experiences, lack of language assessment education due to being a graduate of other related fields, high workloads, and feeling less confident in assessment practices have been reported as reasons for the low level of language assessment literacy (Abrar-ul-Hassan & Douglas, 2020; Goh, 2011; Herrera & Macías, 2015; Hidri, 2021; Lam, 2019; Lan & Fan, 2019; Scarino, 2013; Taylor, 2013; Zhu, 2004).

Teachers' Language Assessment Literacy plays a vital role in leading the way for learners appropriately. LAL training should be created to enhance the quality of language teaching (Herrera & Macías, 2015). As Giraldo (2018) states, "Language teachers need to be able to deliver high-quality assessments for the development of students' language proficiency, which is only possible if they possess the knowledge, skills, and practices of language testing" (p. 241).

2.3.1. Teachers' Language Assessment Literacy and Training

Language teachers require constant training both during their preservice and inservice stages, as the field and the needs of learners are constantly changing. International studies have also emphasized the importance of training on language assessment knowledge (Arias & Maturana, 2005; Cheng, Rogers & Hu, 2004; Diaz, Alarcon & Ortiz, 2012; Frodden, Restrepo & Maturana, 2004; Lopez & Bernal, 2009; Munoz, Palacio & Escobar, 2012; Volante & Fazio, 2007). Alonzo, Leverett & Obsioma (2021), highlighted the interconnected nature of teaching and assessment practices and the need for language assessment literacy training.

Preservice language teachers have expressed their need for practicing assessment knowledge in various studies, and they emphasize the importance of practice in classroom-like situations to increase their competence as they enter the classroom (Kavaklı & Arslan, 2019; Sarıyıldız, 2018). Inservice language teachers also stress their need for training on alternative assessment methods, such as portfolios, integrated language skills, and aspects of culture (Semiz & Odabaş, 2016). Another study by Ölmezer-Öztürk & Aydın (2019) supports this finding and shows inservice language teachers' need for assessment, specifically for listening skills.

Öztürk and Aydın (2018) carried out a study to explore both the general and specific skill-based language assessment literacy of English as a Foreign Language (EFL) instructors at university preparatory schools. They also examined how demographic characteristics might influence this literacy. To gather data, they surveyed 542 EFL instructors using the Language Assessment Knowledge Scale (LAKS), a tool they developed. Their findings revealed that these instructors were most proficient in assessing reading skills, but less so in evaluating listening skills. Interestingly, the study determined that demographic factors generally did not impact the instructors' language assessment literacy, with the exception of those who were members of a testing office. The study highlighted the ongoing need for language teachers to receive regular training in language assessment literacy. This requirement is particularly pressing given the ever-evolving nature of the field and learner needs. This necessity for continuous professional development in language assessment is further underscored by similar findings in international research focused on language assessment training needs. In a study by Yan, Zhang, and Fan (2018) involving three Chinese secondary-level teachers, it was found that they require training in language assessment practices. In a more recent study, Kremmel and Harding (2020) investigated the language assessment literacy requests of a diverse group of 1086 individuals. This group included language teachers, language test developers, and language testing researchers. They analyzed the responses gathered to offer research-backed insights into the training necessities of the participants. Upon the introduction of new assessment reform, Froozi et al., (2019) carried out a study to understand assessment related needs of Iranian EFL teachers. Document analysis and in-depth interviews were also used to develop better understanding on training needs of teachers. In China, Lan and Fan (2019) conducted a study on the training needs of EFL teachers. The study involved 344 EFL teachers who expressed their desire for training to understand concepts of classroom-based assessment and practical knowledge on assessment. Sun and Zhang (2022) conducted a study involving 272 EFL instructors from 146 universities to determine their current levels of language assessment literacy and training needs related to language assessment. The results showed a low level of involvement in language assessment activities and unsatisfactory self-perception of language assessment literacy, indicating a need for training in all aspects of language assessment. Mede and Atay (2017) conducted a study to investigate the language assessment knowledge, training needs, and attitudes towards assessment practices of EFL instructors. An online questionnaire, which was adapted from Vogt

and Tsagari's 2014 study, was used, followed by focus group interviews. The study included 350 EFL instructors as participants. The results revealed a noticeable deficiency in assessment literacy and highlighted a significant demand for training, particularly in areas like test development and feedback. While instructors felt relatively confident in assessing grammar and vocabulary, they acknowledged their lack of proficiency in evaluating productive and receptive skills. The study emphasized the necessity for language teachers to continually enhance and update their knowledge to effectively address instructional challenges.

Despite receiving preservice training on language assessment, language teachers still express a need for more advanced training (Arias & Maturana, 2005; Cheng, Rogers & Hu, 2004; Diaz, Alarcon & Ortiz, 2012; Frodden, Restrepo & Maturana, 2004; Volante & Fazio, 2007; Lopez & Bernal, 2009; Munoz, Palacio & Escobar, 2012). Vogt and Tsagari (2014) also found that language teachers require advanced training for both receptive and productive skills, as well as micro-linguistic skills, despite receiving preservice or in-service training. Similarly, Fulcher (2012) highlights the need for language assessment books that include more real-life assessment activities to further support language teachers.

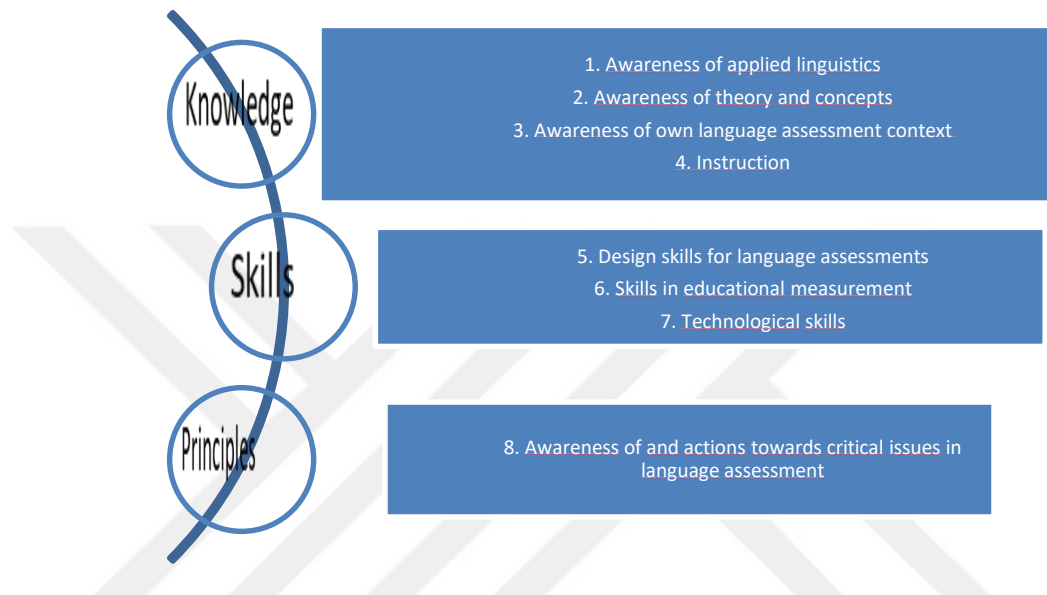
Numerous studies have emphasized the need for in-service training of language teachers. Mahapatra (2016) further supported this conclusion by emphasizing that Web 2.0 tools can enhance the language assessment literacy levels of in-service language teachers. As their assessment literacy levels increased, their assessment practices changed as well. They became more evaluative in their assessment practices, engaged their students more in assessment, and related their assessment with their teaching. They started to learn more about their assessment practices and their students' learning. Various studies have also shown the significance of online courses on language teachers' assessment practices. These studies reveal that language teachers feel more comfortable taking online assessment courses, which contributes to their attitudes towards using assessment terms during their in-class practices (Nier et al., 2013; Riestenberg et al., 2010). In conclusion, the lack of sufficient language assessment education during pre-service training and the feeling of incompetence due to a lack of practical assessment knowledge lead to a demand for language assessment training among teachers. This demand is necessary for language teachers to effectively use assessment tools that can be used to make informed decisions about their students' learning and their own teaching practices. Hill (2017) highlighted

the growing significance of language assessment literacy among language teachers, attributing it to two primary factors. The first is the expanding role of language teachers in student assessment within the classroom. The second factor is a notable shift in assessment practices, moving from a sole focus on assessment of learning to an increased emphasis on assessment for learning. This evolution signifies that assessment in the classroom is no longer limited to evaluating what students have learned, but also involves using assessment as a tool to enhance learning. Hill (2017) pointed out that this paradigm shift has fundamentally altered the objectives of teaching and assessment. Tsagari and Vogt (2017) supported this argument and suggest that these developments necessitate language teachers to be competent in language assessment. They emphasized the need for continuous professional development for language teachers to effectively navigate and implement classroom-based assessment without facing challenges. Mertler (2005) aimed to evaluate and statistically compare the assessment literacy levels of pre-service and in-service teachers. To gather data, he employed a survey tool called "the Classroom Assessment Literacy Inventory," which was designed to measure the assessment literacy in both groups. The findings of this study revealed that in-service teachers possessed a higher level of assessment literacy compared to their pre-service counterparts. Despite this difference, it was noted that both groups exhibited limited knowledge in assessment practices. He concluded that the discrepancy in assessment literacy levels could be attributed to the misalignment between the assessment-related courses offered in pre-service education and the practical needs of teachers regarding assessment practices in real classroom settings.

In summary, it is crucial to provide language teachers with appropriate training to develop their assessment literacy. As Jeong (2013) suggests, teacher training on assessment is necessary to equip teachers with the skills and knowledge to effectively assess their students in the classroom. All pre-service and in-service English language training programs should include opportunities for language assessment literacy development. Giraldo (2018) highlights the importance of language teachers being aware of the eight components of language assessment literacy, which include knowledge, skills, and practices of language testing, in order to enhance their students' language proficiency (see Figure 3):

Figure 3

Eight components of LAL for language teachers



This study's classification was based on a variety of sources, including definitions of language assessment literacy, language assessment course books, the items proposed in the PKAQ, and general competency indicators recommended by MoNE (2017).

Table 1

Theoretical framework for language assessment literacy

Psychometric Features of Test	Basics of Test Development	Assessing Language Skills	Functions of Assessment
Item Analysis (IF,ID,CD)	Procedures in Language Test Design	Assessing Reading Comprehension	Formative Assessment
Reliability of the Test	Deciding on What to Test	Assessing Listening	Summative Assessment
Validity of The Test	Writing Test Specifications/Blueprint	Assessing Speaking	Diagnostic Assessment
Interpreting Test Scores	Selecting Items for the Test	Assessing Writing	Peer and Self Assessment

Knowledge of Statistics	Preparing Items	Performance Assessment
	Test Administration	Portfolio Assessment
	Giving Feedback to students	
	Developing Fair and Objective Scoring Rubric	
	Ethical Considerations in testing	

2.3.2. General Competencies of Teaching Profession in Turkey

One of the basis of the theoretical framework for the study is the General Competencies of Teaching Profession suggested by MoNE (2008). The components of competencies , in fact, are the milestones for the teaching profession that shed light on teachers' way to enhance their teaching.

Competency refers to the level of ability to perform a particular task or skill. Teacher competency, on the other hand, is defined as the degree of specific skills and knowledge required to effectively carry out their professional duties (MoNE, 2008). In accordance with the Basic Law of the Ministry of National Education, which emphasizes the importance of identifying professional competencies, a commission was established in 1999 comprising representatives of MoNE and various universities. The commission proposed the following competency domains:

- Teaching and Education Competencies
- General Information and Cultural Competencies
- Special Areas Knowledge and Skills Competencies

In 2006, the Directorate General for Teacher Training and Development organized a workshop that played a pivotal role in refining the general competencies for educators. This collaborative effort led to the identification of six major competency domains, which were then meticulously

broken down into 39 subdomains, encompassing a total of 244 performance indicators, as reported by the Ministry of National Education. These domains included Personal and Professional Values, Knowledge about Students, the Teaching and Learning Process, Monitoring and Evaluating the Teaching and Development Process, Relations with School, Parents, and Society, as well as Curriculum and Content Knowledge. Each of these areas was carefully delineated to provide a comprehensive framework for teacher competencies, aiming to enhance the quality and effectiveness of education. Based on the competency domains listed above, MoNE developed 14 subject-specific competencies for English Language Teachers at the secondary level in 2008. These competencies include:

- Using effective communication skills in English language teaching
- Designing and implementing lesson plans that promote student learning
- Choosing appropriate materials and resources for language teaching
- Creating and implementing effective assessment strategies
- Promoting learner autonomy and responsibility
- Providing effective feedback to students on their performance
- Managing classroom dynamics effectively
- Creating and maintaining a positive and supportive learning environment
- Using technology effectively in language teaching
- Demonstrating knowledge of the principles of language acquisition and language learning
- Demonstrating knowledge of the English language and its structure
- Promoting cross-cultural understanding and awareness
- Encouraging students to develop critical thinking skills
- Engaging in continuous professional development and growth.

Regarding the competency areas listed above, subject specific competency areas for English Language Teachers at secondary level include Planning and Organizing the Teaching Process of English, which entails strategizing and structuring English lessons; Developing Language Skills, focused on enhancing students' English language abilities; Monitoring and Evaluating Language Development, to track and assess students' progress in English; Collaboration with

Parents and Society, which involves engaging with students' families and the wider community; and Maintaining Professional Development, emphasizing the continuous improvement and updating of the teachers' own skills and knowledge in the field.

In 2008, Turkey developed its own national competency framework in compliance with the European Union Competency Framework. The Competencies Framework for Turkey was designed to align with the European Competencies Framework for Higher Education, the European Higher Education Competency Framework, and the Higher Education General Competencies Framework in the context of the Bologna Process and Lisbon Strategy. The Higher Education Competency Framework and Higher Education General Competencies Framework were implemented in 2011 under this framework (MoNE, 2017).

The educational world is constantly changing and developing, and it is important for teachers to keep up with these changes. Therefore, there was a need to update the General Competencies for Teaching Profession. The current model includes three Competency Domains with 11 subdomains and 65 indicators (Figure 4), which were revisited and updated to ensure that teachers have the necessary skills and knowledge to pursue their profession effectively and efficiently.

Figure 4

General competencies for teaching profession (MoNE, 2017)

A Professional Knowledge	B Professional Skills	C Attitudes and Values
A1. Content Knowledge	B1. Planning of Education and Teaching	C1. National, Moral and Universal Values
She/he has an advanced and critical perspective on theoretical, methodological and factual knowledge in his/her subject field.	She/he plans education and teaching processes effectively.	She/he observes national, moral and universal values.
A2. Pedagogical Content Knowledge	B2. Creating Learning Environments	C2. Approach to Students
She/he has a good knowledge of the curriculum and pedagogical content knowledge of her/his subject area.	She/he prepares appropriate teaching materials and builds an healthy and safe learning environments, where effective learning can be achieved for all students.	She/he has an attitude that supports the development of students.
A3. Knowledge on Legislation	B3. Managing the Teaching and Learning Process	C3. Communication and Cooperation
As an individual and teacher, she/he conducts her/himself according to the legislation related to her/his duties, rights and responsibilities.	She/he manages the teaching and learning process effectively.	She/he establishes an effective communication and cooperation with students, colleagues, families, and other educational stakeholders.
	B4. Assessment and Evaluation	C4. Personal and Professional Development
	She/he uses the methods, techniques and tools of assessment and evaluation that fit for purpose.	By carrying out self appraisal she/he participates in personal and professional development activities.

However, with the adoption of the updated version of General Competencies for Teaching Profession and the addition of a general competency related to "C4: Personal and Professional Development," the need for Subject Specific Competencies for Teaching Profession was eliminated. For this study, the Competency Area "B4: Assessment and Evaluation" will serve as the basis for examining changes before and after the online workshop, which will be designed based on the results of the needs analysis. The competency indicators to be considered under this competency area are listed below in Figure 5.

Figure 5

General competencies for assessment and evaluation (MoNE, 2017)

B	COMPETENCY DOMAIN: PROFESSIONAL SKILLS
	SCOPE: This competency domain covers the teacher's competencies of educational planning, creating learning environments, managing the teaching and learning process, and monitoring and evaluation regarding the his/her in-class and out-of-class practices.

COMPETENCIES	COMPETENCE INDICATORS
B4. ASSESSMENT AND EVALUATION She/he uses the methods, techniques and tools of assessment and evaluation that fit for purpose.	B4.1. She/he prepares and uses assessment and evaluation tools suitable to his/her subject area and stages of growth and development of students.
	B4.2. She/he makes use of both formative and summative assessment methods.
	B4.3. She/he carries out an objective and fair assessment and evaluation.
	B4.4. She/he provides accurate and constructive feedbacks to students and other stakeholders in accordance with results from assessment and evaluation.
	B4.5. She/he re-arranges the teaching and learning processes in accordance with the assessment and evaluation results.

In this figure the competence indicators of the assessment and evaluation competence is listed in detail. It is underlined that competence in assessment and evaluation cover preparing and applying the appropriate assessment methods; use of various assessment methods, being fair and the use of the assessment results regarding both his/her teaching practices and students' learning processes.

2.4. Teachers' Language Assessment Literacy Perceptions

Various studies, including those by Rogers, Chan, and Yu (2007), Shohamy et al. (2008), and Munoz, Palacio, and Escobar (2012), have highlighted the importance of using appropriate assessment methods to improve teaching and enhance students' learning. However, research also suggests that teachers tend to prefer traditional assessment methods and do not always use assessment results to inform their teaching practices. This discrepancy between teachers' beliefs and practices can be attributed to factors such as class size, time constraints, standardized tests, workload, teaching context, and lack of training (Chan, 2008; Rogers et al., 2007; Shohamy et

al., 2008). Additionally, experience plays a significant role in the assessment practices of language teachers, as noted by Hakim (2015), who suggests that more experienced teachers tend to use assessment concepts more effectively in their practices.

The assessment methods used by EFL instructors vary depending on the classroom level. A study conducted by Zhang and Burry-Stock (2003) with 297 teachers aimed to identify the perceptions of assessment skills and practices among EFL instructors. The study findings revealed that elementary level teachers prefer using alternative assessment methods, while secondary teachers prefer more objective methods. Additionally, the study showed that teachers across content areas preferred using a variety of assessment techniques.

A number of studies have explored EFL instructors' perceptions and practices of assessment. Zhang and Burry-Stock (2003) surveyed 297 teachers to investigate their assessment practices, and found that elementary-level teachers tended to prefer alternative assessment methods while secondary-level teachers favored more objective methods. Across all content areas, teachers reported using a variety of assessment techniques. In another study, Munoz et al. (2012) conducted surveys, interviews, and written reports of experiences with 62 language teachers at a private university in Colombia, and found that although teachers believed in the value of using assessment results to improve teaching, they did not consistently apply this belief in their assessment practices. The researchers recommended greater use of formative assessment practices to facilitate the use of assessment results for instructional improvement. Similarly, Jannati (2015) conducted semi-structured interviews with 18 Iranian EFL instructors to explore their perceptions and in-class assessment practices, finding that year of teaching did not significantly affect teachers' assessment preferences and that teachers did not consistently apply their assessment knowledge in practice. Finally, Hakim (2015) used a questionnaire to examine language teachers' perceptions, practices, and the impact of teaching experience on their assessment knowledge, and found that while all teachers had some basic knowledge of language assessment techniques, more experienced teachers tended to use in-class practices more effectively than novices.

Fulcher's (2012) LAL Survey was adapted by Janatifar and Marandi (2018) to define language assessment literacy components in language teachers' perceptions. The questionnaire

resulted in four components: test design and development, large-scale standardized testing, classroom assessment, and validity and reliability. The importance of hands-on activities and skill-based approaches were also emphasized. However, despite positive attitudes towards alternative assessment methods, in-service language teachers struggle to put their knowledge into practice due to heavy workloads, large class sizes, and extracurricular activities (Büyükkarcı, 2014). Additionally, Gönen and Akbarov (2015) concluded that centralized assessment systems, syllabi, and students' educational background impact in-service teachers' assessment practices, despite their positive beliefs.

While Aksu & Ataç (2012) suggest that the lack of sufficient assessment knowledge is a key factor in in-service teachers' less favorable beliefs about assessment, other studies indicate that teachers often maintain positive beliefs and attitudes towards assessment regardless of the inadequacies in pre-service assessment training courses, as evidenced by research from Büyükkarcı (2014), Han & Kaya (2014), and Öz (2014). Additionally, Şahin (2019) discovered that pre-service EFL teachers generally found their Language Assessment Literacy training in the English Language Teacher Education Curriculum (ELTEC) to be satisfactory. However, they felt less confident in areas such as alternative assessment tools, formative assessment tools, and the assessment of productive language skills. Öz and Atay (2017) conducted a study to reveal language teachers' perceptions and practices of language assessment. Twelve EFL instructors with teaching experience ranging from 1-15 years participated in the study through semi-structured interviews. The findings showed that despite adequate language assessment knowledge, teachers cannot efficiently put this knowledge into practice. It was also revealed that teaching experience has no significant effect on assessment literacy.

Although research shows diversity in teachers' beliefs and perceptions regarding language assessment practices, one thing remains constant: the effect of these beliefs and perceptions on their practices. Firoozi et al. (2019) emphasize the significance of teachers' assessment-related perceptions when adopting new assessment policies. Building on these findings, Alonzo et al. (2021) stress the effect of changes in assessment policies on teachers' beliefs and conceptions, which ultimately enhance teaching practices. As teachers understand the connection between teaching and assessment, they are more motivated to apply and adopt changes in policies. Thus, studies on language assessment literacy perceptions and beliefs are fundamental to

understanding the underlying reasons for teachers' assessment practices. If change is planned, teachers' beliefs and perceptions should be considered.

2.5. Empirical Studies on Language Assessment Literacy

Language Assessment Literacy (LAL) is a relatively new term that researchers across various fields mentioned about. A thorough examination of existing literature reveals numerous studies centered on different aspects of LAL. These studies primarily explore areas such as the LAL-related needs of pre-service and in-service teachers, the development of LAL among these groups, and their perceptions and beliefs concerning LAL competence, as well as its application in classroom practices. For instance, Firoozi et al. (2019) conducted a study using semi-structured interviews to delve into the assessment literacy of Iranian EFL instructors and how it relates to the classroom assessment needs and assessment policies in Iran. The study involved 15 Iranian EFL instructors and revealed that these teachers generally lacked adequate assessment knowledge and faced challenges in developing tests. Despite Iranian assessment policies permitting alternative assessment methods like portfolios or self-assessment, the teachers predominantly followed traditional assessment techniques, indicating a gap between the official assessment policies and the actual classroom practices of EFL teachers. Similarly, Sultana (2019) research focused on understanding the perceptions of Bangaldeshian ELT teachers and how these perceptions are reflected in their classroom practices. This study, which involved semi-structured interviews with 10 teachers from various schools highlighted a shortfall in assessment training for in-service teachers, despite their responsibilities in test development. Furthermore, the study underlined the preference of teachers for traditional assessment methods over more contemporary approaches.

Empirical studies on language assessment literacy in Turkey have also emphasized the inadequacy of pre-service language assessment courses. Student teachers' beliefs on assessment courses were often found to be incomplete, attributed either to their beliefs or the limited time spent on assessment courses during pre-service education. Erçetin (2016) identified time constraints as a factor and stressed the insufficiency of a single course to cover all issues or concepts of assessment. Furthermore, the Turkish education system heavily relies on summative

assessment, which discourages teachers from developing themselves in alternative assessment methods (Gönen & Akbarov, 2015; Şahin, 2019).

Similar to international studies, pre-service teachers in Turkey stressed the importance of theoretical knowledge on language assessment, but they lacked opportunities to apply this knowledge in practice. Yastıbaş (2018) found that in-service language teachers developed their professional skills through experience, and the study revealed that participants graduating from education faculties were content with their pre-service assessment training, while participants graduating from other faculties deemed the training ineffective.

Two studies that highlight the insufficient practical application of theoretical knowledge in language assessment were conducted by Hatipoğlu (2010) and Erçetin (2016). Both studies emphasized that having only one course during preservice training is not enough to cover all the issues and concepts of language assessment, and that there is not enough time to practice the theory. Sevime-Şahin (2019) investigated the effectiveness of the English Language Testing and Evaluation course on preservice training, with 178 participants. The study revealed that the participants mainly concentrated on aspects such as summative assessment, diagnostic and achievement tests, and the foundational knowledge of assessment literacy. The participants noted a deficiency in practical training, integrated language testing, as well as skills in interpretation and evaluation. They also identified a gap in understanding alternative assessment methods and the application of test results. Consequently, there is a clear need for a revision of the ELTE course content to incorporate more hands-on training, integrated language testing experiences, and enhanced interpretation and evaluation skills. However, in contrast to the previous studies that emphasize the theoretical nature of preservice language assessment courses, Yetkin (2015) argues that student teachers can practice what they have learned through assignments and school practicum courses. His study aimed to investigate the perceptions of 30 candidate English teachers on their perceptions of the effect of assessment course on the development of assessment literacy utilization of assessment strategies in primary/junior settings in a state university. Mede and Atay (2017) also highlight the need for more training on language assessment in the Turkish context. They suggest that the training offered to language teachers should include various forms of assessment methods in language preparatory schools. In another study conducted by Hatipoğlu (2015), which involved 124 preservice English

language teachers, the expectations of participants with regard to English Language Testing and Evaluation (ELTE) courses were investigated using questionnaires and interviews. The results showed that most of the participants believed that the ELTE course they took during undergraduate education would be adequate for all their assessment practices while teaching. However, despite their four-year education, student teachers still had inadequate assessment-related knowledge, which would reflect on their classroom practices as language teachers. Therefore, greater importance should be given to testing and assessment courses taken throughout undergraduate education. In her dissertation, Sarıyıldız Canlı (2023) stressed the inservice training need regarding Language Assessment Literacy of the teachers working both formal and non formal institutions. While quantitative part of the study was carried out with 300 inservice teachers to define the language assessment related needs of teacher, 20 inservice teachers took part in interviews designed to explore assessment practices of Language Teachers. The study also investigated the effects of the factors such as kind of institution, grade level, level of experience and education. The current research study aimed to examine the LAL levels and training needs of in-service EFL teachers who work in different formal and non-formal education institutions. The results of the study also put emphasize on the significance of the mastery of the Language Assessment Literacy levels on instructional improvement. Echoing the results of the above mentioned study, a recent MA thesis by Doğan (2023) revealed that English Language Teachers working at K12 schools in Turkey are lack of sufficient assessment related knowledge in need of training.

2.6. Chapter Summary

Studies with preservice language teachers stress the efficacy of language assessment courses, but they also put emphasis on their need of practicing the language assessment related theoretical knowledge. (Arias & Maturana, 2005; Cheng & Rogers & Hu, 2004; Diaz, Alarcon& Ortiz, 2012; Frodden, Restrepo & Maturana, 2004; Lopez & Bernal, 2009; Munoz, Palacio& Escobar, 2012; Volante & Fazio, 2007). Graduating without gaining enough experience in practicing language assessment related knowledge, teachers start with incompetent knowledge base in using appropriate assessment methods in order to make valid and reliable educational

decisions about both their teaching and learning (Jin, 2010; Jeong, 2013; Lam, 2015). Having low level of practical language assessment literacy, language teachers tend to use traditional methods and familiar methods that they feel more competent about. Thus, this may lead teachers to make inappropriate decisions regarding the choice of assessment tools and educational decisions (Kiomrs, Abdulmahdi & Naser, 2011).

Upon examining the literature related to language assessment literacy, it is very apparent that the need for training and the feeling of incompetency of the language teachers in the field starts just from the preservice education stage of the language teachers. Many studies stress the gap between theory and practice of the knowledge. Teachers emphasize that they gain theoretical knowledge but feel incompetent in practicing the knowledge. As a result, teachers cannot put theory into action efficiently. Inefficiency in putting the theoretical knowledge into action causes either the use of appropriate assessment methods to make educational decisions or the frequent use of familiar methods of which results does not serve their aim to improve their own teaching and students' learning. Using inappropriate method of assessment due to the lack of practical knowledge of language assessment may have a life-long impact on one's career due to the educational decisions that are taken as a result of the assessment practice.

On the contrary to this finding, having enough knowledge of language assessment, language teachers are more able to use assessment as a tool to improve their own teaching and choose appropriate assessment method for their purpose of assessment (Herrera & Marcias, 2015). Chan (2008) indicated that as long as teachers are not affected with overload and time consuming activities; teachers prefer using alternative methods and use assessment to enhance their own teaching. Therefore, teachers may have enough language assessment related knowledge base, but they cannot perform their knowledge due to overload and time consuming activities; teachers prefer using alternative methods and use assessment to enhance their own teaching.

To conclude, it is apparent that level of language assessment literacy has a great impact on in-service Language Teachers' assessment practices. The tendency to use alternative assessment methods besides traditional assessment methods is strictly related to the content of

language assessment courses. It is worth noting that extracurricular works has a role in language teachers' assessment preference.



3. METHODOLOGY

This dissertation used a case study design to assess EFL teachers' current assessment literacy levels (knowledge, perceptions, and competencies) identify their assessment needs and design an online workshop for improving language assessment literacy competencies of EFL teachers. This chapter provides an overview of the methodology adopted for the study. It includes a description of the pilot study stage, the participants, research methodology, data collection instruments, data collection procedures, research questions, and data analysis techniques.

The study aims to address the following research questions:

1. What are the language assessment literacy needs of EFL in-service teachers?
2. What are the competency indicators that EFL in-service teachers have before the language assessment literacy (LAL) workshop?
3. What competency indicators do EFL in-service teachers acquire through LAL workshop?
4. Do LAL workshop effectively improve the language assessment literacy competencies of EFL in-service teachers?

3.1. Setting

The study was conducted in Mersin, Turkey, and targeted English Language Teachers working at the secondary level in four central districts of the Mersin. The population of the study comprised of 524 English Language Teachers. The English language teaching hours for 5th and 6th graders is three hours per week, while 7th and 8th graders receive four hours per week. In addition, elective English courses ranging from two to seven hours per week are offered according to school schedules. All English Language Teachers in the study followed the English language teaching Curriculum recommended by the Ministry of National Education (MoNE).

3.2. Participants

The target population of this study comprised of all English language teachers working at state secondary schools in Mersin, Turkey. However, due to the voluntary nature of participation, the number of volunteers varied for different stages of data collection. The study consisted of two groups of participants (See Appendix A for the consent form).

The first group consisted of participants who completed the Practical Knowledge about Language Assessment Questionnaire (PKAQ), which aimed to gather information about the language assessment-related needs of language teachers. The second group comprised of six volunteered teachers who participated in interviews, online language assessment literacy workshop, and pre- and post-tests (PKAT) as part of the study (See Appendix B for the ethics report).

During the recruitment process, the researcher used various methods such as social media groups for English Language Teachers in Mersin (Facebook, Instagram), attending three meetings held by the Provincial Directorate of National Education, and receiving support from the Assessment and Evaluation Department of Mersin.

3.2.1. Participants of the Pilot Study for PKAQ

Two pilot studies were conducted for this study. The first pilot study was conducted on May 29, 2018, where 225 English language teachers participated in a meeting. However, the results obtained were not valid and reliable, leading to the conduct of a second pilot study. The second pilot study involved 82 ELT teachers who volunteered to participate in the study. The age range of the participants was between 22 and 55 years, with at least one year of teaching experience. The teaching experience of the participants ranged from 1 year to 25 years or more. All participants were graduates of ELT, Literature, or related fields.

3.2.2. Participants of PKAQ

After obtaining the necessary permissions from the Provincial Directorate of National Education in Mersin, the researcher presented the study at a meeting for the heads of English Language Teachers of secondary level schools on December 7, 2021. During the meeting, the study was briefly explained and some of the attendees volunteered to fill out the forms immediately. However, due to pandemic precautions, some of the participating teachers declined to fill out the forms and requested that the PKAQ be transferred to an online Google form. In total, 101 teachers completed the forms, with ages ranging from 26 to 46 and above. All participants had at least one year of teaching experience.

3.2.3. Participants of the Online Workshop, Interviews and PKAT

Although there is a large number of ELT teachers working at the secondary level, only nine EFL teachers volunteered for the second part of the study. Prior to consenting to participate in the study, they wanted to ensure that the study would not require them to take on any additional work in their classes. Specifically, they made it clear that they would only participate if they were not required to do anything extra in their classes. Due to time constraints, the researcher had to proceed with a limited number of volunteers. To provide a detailed explanation to the volunteer teachers, the researcher organized a Zoom meeting on February 17, 2022. During the meeting, the steps and purpose of the study were explained, and the teachers were asked to provide more details on their language assessment literacy needs. After the meeting, the teachers were asked to complete the PKAT. However, two of the volunteers chose not to continue participating in the study and did not provide reasons for their decision. During the third week of the online workshop, one of the volunteer teachers had to stop participating due to a sudden family health issue. As a result, a total of six volunteer teachers from five different schools, all of whom were female, participated in the study. The participants had between 8 and 22 years of teaching experience.

3.3. Research Design

This study adopts a case study research design. Bogdan and Biklen (1998, p.59) define a case study as "a detailed examination of one setting, or a single subject, a single depository of documents, or one particular event." Mackey and Gass (2022, p. 308) explain that case studies aim to provide a holistic description of learning or use within a specific population and setting. However, it is essential for researchers to exercise caution when generalizing results from case studies, particularly due to the limited number of participants. Mackey and Gass (2022, p. 311) recommend that generalizations from individual or small groups to larger populations should be made tentatively and with extreme caution. Johnson (1993), cited in Mackey and Gass (2022, p. 310), emphasizes the advantage of case studies in explaining "the complexities of particular cases in specific contexts over quantitative research designs" such as correlational, survey, and experimental research.

Given the study's focus on understanding the language assessment-related needs of language learners and the intent to use both quantitative and qualitative data collection methods, employing a case study design is justified for several reasons:

- **In-depth Analysis:** The case study design allows for an in-depth exploration of the specific needs and experiences of language learners in relation to language assessment.
- **Contextual Understanding:** Language learning and assessment are heavily influenced by context - including cultural, educational, and individual factors. A case study approach enables the examination of these contextual factors in detail, providing a richer understanding than what might be obtained through a more generalized survey or experimental study.
- **Combining Data Types:** The flexibility of case studies in accommodating both qualitative and quantitative data is ideal for this study. Qualitative data can capture the detailed experiences and perspectives of learners, while quantitative data can provide measurable insights into their proficiency and progress.

- **Limited Number of Participants:** Case studies are well-suited for studies with a smaller sample size. They can provide comprehensive insights into each participant's language assessment needs, which might be lost in a larger-scale study.
- **Exploratory Nature:** If the study aims to explore new or under-researched areas within language assessment, the case study design is particularly suitable. It allows for an exploratory approach where emerging themes and patterns can be identified and investigated in depth.
- **Practicality and Resource Efficiency:** With a limited number of participants, a case study approach is often more practical and resource-efficient, allowing for a thorough exploration of each case within the constraints of the study.

In summary, a case study design is justified for this study as it aligns well with the goals of gaining a deep, contextual, and nuanced understanding of the language assessment needs of a specific group of learners, utilizing a combination of data collection methods within a manageable scope.

3.4. Instruments

For the first part of the study, which was the needs analysis, the data were obtained from several sources, including the Professional Knowledge of Assessment Questionnaire, originally developed by Fulcher (2012) and adopted by Farhady & Tavassoli (2018), the Professional Knowledge of Assessment Test (Farhady, 2016), which was used as a pre-test, and the General Competency Indicators for English Language Teachers suggested by the Ministry of National Education. Additionally, a focus group interview was conducted with English language teachers who volunteered to participate in the second part of the study.

Regarding the second part of the study, the pre- and post-professional Knowledge of Assessment Test scores, interviews, and reflective notes were utilized to evaluate the effectiveness of the online Language Assessment Literacy workshop and to understand its impact on the Language Assessment literacy competencies of the participating teachers.

3.4.1. Professional Knowledge of Assessment Questionnaire (PKAQ)

The main instrument utilized for this study is the Professional Knowledge of Assessment Questionnaire (PKAQ), which was originally developed by Fulcher (2012) and adopted by Farhady & Tavassoli (2018). An adapted version of the questionnaire was used after the completion of the piloting stage. The original survey consisted of 12 questions on biodata, one closed response question with 24 sub-questions on assessment topics. However, the adapted version of the questionnaire consists of 13 items (see Appendix C).

Items 1-7 aimed to gather demographic information about the participants, such as gender, age, teaching experience, educational background, and teaching position. Item 8 was used to check whether the participant had taken a language testing course at the university or not. Item 9 was an open-ended question to obtain the participant's opinion on language assessment related needs. Items 10-12 were in the form of a 4-point Likert scale. Item 10 asked the participants about the importance of language assessment related topics, Item 11 inquired about the participants' knowledge and understanding of language assessment related topics listed in the questionnaire, Item 12 aimed to gather information on the extent to which participants apply their language assessment related knowledge in their classes, and the last item was used to obtain comments from participants regarding their language assessment related needs.

The overall objective of using PKAQ was to establish a solid foundation for designing an online language assessment literacy workshop that would have an impact on the language assessment related competencies of the participants.

3.4.2. Professional Knowledge of Assessment Test (PKAT)

To assess the effectiveness of the workshop, the Professional Knowledge of Assessment Test (PKAT) developed by Farhady (2016) was used as both a pre-test and a post-test in this study (see Appendix D). The PKAT consists of 35 items, which were checked for alignment with the components of the theoretical framework before administering the test (see Appendix E for the alignment table). Experts working at the Assessment and Evaluation Center of MoNE

also reviewed the match between the items of PKAT and the components of the theoretical framework.

The items in PKAT can be grouped as follows: items 1-5 require participants to match situations with appropriate test types, items 6-10 ask participants to order the stages of test design, items 11-15 require participants to match scenarios with listed statistical techniques, items 16-23 require participants to complete sentences regarding psychometric features of testing, items 24-28 require participants to match classroom assessment activities with corresponding assessment techniques, and items 29-35 require test-takers to choose the correct options regarding language skills assessment. The following tables summarize the items and their related components of the theoretical framework for language assessment literacy:

Table 2

Distribution of PKAT items and language assessment literacy topic

PKAT Item	Language Assessment Literacy Topic
1-5	Test Typology and Purposes
6-10	Test Development and Design
11-15	Test Analysis and Use
16-23	Psychometric Features of Testing
24-28	Classroom Assessment Practices
29-35	Language Skills Assessment

Table 2 above indicates the distribution of the items with the Language Assessment Literacy topics. Also, the alignment between PKAT Items and Theoretical Framework of LAL components (See Appendix F) had been ensured priori to the use of PKAT (See Table 3).

Table 3

Alignment between the PKAT items and the theoretical framework for LAL

Items	Related Components of Theoretical Framework
1-5,19,22,29	Functions of Assessment
6-10,16-18,20,24-28	Basics of Assessment
11-15,21,23	Psychometric Features of Testing
30-35	Assessing Language Skills

In the proposed framework of the study, the PKAT items were aligned with distinct components of the theoretical framework for LAL. Items 1-5, 19, 22, and 29 are linked to the Functions of Assessment, shedding light on the overarching goals and roles of language assessment in educational contexts. Conversely, items 6-10, 16-18, 20, 24-28 examine the Basics of Assessment, summarizing the fundamental principles and concepts essential to the construction and understanding of assessment tools. A critical examination of Psychometric Features of Testing was represented by items 11-15, 21, and 23, emphasizing the statistical validity essential in testing methodologies. Lastly, items 30-35 are for Assessing Language Skills, focusing on the practical aspects of evaluating specific language competencies such as reading, writing, speaking, and listening.

3.4.3 Semi-structured Interviews

The study utilized interviews as a primary source of qualitative data. Interviews are a valuable data collection method as they provide an in-depth understanding of people's ideas, opinions, and understanding (Yeo, Legard, Keegan, Ward, Nicholls & Lewis, 2013). The types of interviews utilized depend on how the interview structure and content are designed (Bryman, 2008). There are four types of interviews: informal, retrospective, structured and semi-structured (Ary, Jacobs and Razavieh, 1996). Informal interviews are less formal than structured and semi-structured interviews, and allow for participants to share their thoughts on the issue

being explored. Retrospective interviews require respondents to remember and reconstruct past events. In a structured interview, the interviewer uses a predetermined set of closed questions in a fixed order. In contrast, in a semi-structured interview, the researcher uses a set of pre-prepared questions that can be flexible in terms of what and how to ask, and the interviewer has some discretion to ask additional questions based on the responses given (Bryman, 2008).

The use of semi-structured interviews was appropriate for this study, as it allowed for a detailed exploration of participants' backgrounds and experiences related to language assessment literacy, as well as their educational contexts (Kruger & Casey, 2000). Since semi-structured interviews allow for a certain degree of flexibility in the interview process, which can be beneficial in exploring the depth and complexity of participants' experiences, perceptions, and practices, semi-structured interviews were used to gain a deeper understanding of language teachers' assessment-related needs, and participants were asked to share their needs based on the components of the language assessment literacy framework. They were asked three questions after completing the first half of the workshop to find out their opinions on the course content (See Appendix G for the suggestions). Participants were also asked to reflect on their in-class experiences and share what they had learned after each lesson. In the middle of the six-week workshop, volunteers were invited to a café for an informal interview to discuss their feelings, hesitations, and awareness raised on the significance and systematic nature of the assessment procedure.

3.4.4 Reflective Notes

The participants were asked to reflect on their current assessment practices, the challenges they face during their assessment practices, and the new assessment practices they learned from the workshop regarding language assessment literacy. After each session, participants were asked to write reflective notes and send them to the researcher via email (see Appendix H for the prompts). Regarding the psychometric features of the test, participants were presented with five caricatures and asked to reflect on the cases depicted in the visuals.

3.4.5 General Competency Indicators for English Language Teachers

In this study, Competency Area "B4: Assessment and Evaluation" (See Figure 6) was used as the base in order to understand language assessment related needs of in-service language teachers, the change before and after the online workshop that had been designed according to the results of the needs analysis.

Figure 6

General competencies of teaching profession: Assessment and evaluation (MoNE, 2017)

B	COMPETENCY DOMAIN: PROFESSIONAL SKILLS
	SCOPE: This competency domain covers the teacher's competencies of educational planning, creating learning environments, managing the teaching and learning process, and monitoring and evaluation regarding the his/her in-class and out-of-class practices.
COMPETENCIES	COMPETENCE INDICATORS
B4. ASSESSMENT AND EVALUATION She/he uses the methods, techniques and tools of assessment and evaluation that fit for purpose.	B4.1. She/he prepares and uses assessment and evaluation tools suitable to his/her subject area and stages of growth and development of students.
	B4.2. She/he makes use of both formative and summative assessment methods.
	B4.3. She/he carries out an objective and fair assessment and evaluation.
	B4.4. She/he provides accurate and constructive feedbacks to students and other stakeholders in accordance with results from assessment and evaluation.
	B4.5. She/he re-arranges the teaching and learning processes in accordance with the assessment and evaluation results.

Both of the quantitative data collection instruments had been aligned with the competency indicators since the language assessment related needs of in-service teachers and the effectiveness of the online workshop on the language assessment skills and competencies of in-service language teachers cannot be defined and evaluated separately.

3.5. Data Collection Procedure

This study employs a comprehensive data collection procedure that encompasses a series of structured phases, each tailored to meticulously gather information for the research. In the

following section, the three key phases of the data collection process will be outlined: the pre-workshop, workshop, and post-workshop phases. Each of these stages is critical for assessing the effectiveness of the language literacy workshop on the participants.

3.5.1. Language Assessment Literacy Workshop

In the field of language assessment literacy, interventions/treatments are designed to engage specific groups of teachers, pre-service or in-service, who have been selected based on identified needs, with the aim of improving their assessment practices. Numerous studies have employed such interventions/treatments to explore their language assessment literacy and determine whether they increase teachers' awareness and lead to seeking and empowering further practices for their professional development. Language assessment courses or workshops can be highly beneficial to participants. Not only can they gain a better understanding of the core principles of language assessment, but they can also apply the new knowledge and technical considerations to empower their assessment practices in the classroom. To achieve this, the researcher developed a workshop for EFL teachers in Mersin, Turkey, aimed at exploring and improving their competencies and, as well as identifying their experiences in assessment in their language classrooms (see Appendix I for the workshop syllabus).

In line with the research aims of the present study, this workshop aimed to evaluate how teachers develop their language assessment literacy skills and competencies. The analysis was carried out in four key areas:

1. Understanding the language assessment related needs of in-service language teachers and utilize the language assessment competencies and theoretical framework of the study to develop workshop syllabus
2. Measuring participants' assessment knowledge before and after the workshop, as well as assessing the quality of the workshop itself.
3. Gathering their perceptions of the workshop and how it related to their current needs, as well as its impact on their language assessment literacy skills, competencies and practices in the language classroom.

4. Tracking participants' development pathways through analysis of reflective notes

Based on the exploratory action research conducted during phase 1 and phase 2 of data collection, it was determined that there was a need for a workshop to be developed. Therefore, the workshop was designed to address their identified needs. The following sections outline the stages involved in the development of the workshop, including the creation of materials, instruction, and evaluation.

Figure 7

Procedure for developing online language assessment literacy workshop



The diagram presented in Figure 7 illustrates the three main phases of the workshop: pre-workshop (development), during the workshop (instruction), and post-workshop (evaluation). During the pre-workshop phase, data from Phase 1 PKAQ and Phase 2 (Interviews) were collected and analyzed to identify the specific needs of EFL teachers related to language assessment literacy. This research triangulation allowed for a comprehensive understanding of the teachers' requirements, which were then used to develop an outline of the workshop, including aspects such as course length, medium of instruction, and materials and content.

Developing Language Assessment Literacy Workshop can be described in three stages. The first one can be summarized as the pre-workshop phase that suggest the steps before the preparation of workshop. The workshop phase covers the steps followed throughout workshop. Post-workshop stage deals with the evaluation of the data to understand the effect of the workshop on language assessment skills and competencies of in-service language teachers.

Pre-Workshop Phase:

All the procedures followed within this phase is related to identifying language assessment related needs of in-service language teachers. Before developing the workshop, the researcher adapted the Professional Knowledge of Assessment Questionnaire (PKAQ) to understand the language assessment needs of the participants. The adapted version of the PKAQ consisted of 13 items. On May 29th, 2018, a meeting was held in a conference room at a school where 225 secondary school EFL teachers were invited to gain information about the study's aim, significance, and timeline. Although all 225 teachers filled out the questionnaire, only 185 provided valid answers to the items, and out of 225, only 35 of them volunteered for the entire study. An email group named #mersineltassessmentsliteracy was created by the researchers. However, the analysis of the data gathered from the pilot study did not yield valid and reliable results, and therefore, the pilot study was repeated.

A second pilot study was conducted to assess the reliability and validity of PKAQ among 88 secondary school language teachers in Mersin. The study included factor analysis to determine its effectiveness. The results of both exploratory and confirmatory factor analysis indicated that the PKAQ was highly reliable and valid for use among the target population.

Various strategies were employed by the researcher to reach out to as many secondary school language teachers as possible, including using social media groups and attending meetings held by the Provincial Directorate of Assessment and Evaluation. A total of 101 language teachers participated in the needs analysis phase of the study, out of which 99 responses were deemed valid for analysis. The collected data were subjected to confirmatory factor analysis (CFA) and descriptive statistics were computed to gain insights into the participants' language assessment needs.

After obtaining the results of PKAQ, focus group interview was conducted on February 17th, 2022 to gain a deeper understanding of the assessment literacy needs of the participants. Based on the findings of the PKAQ and the theoretical framework for language assessment literacy and General Competency Indicators related to Assessment and Evaluation, workshop syllabus was developed. Considering the components of theoretical framework, the workshop duration had been set as 6 week. The topics were as following:

- 1st week: Basics of Testing
- 2nd week: Functions of Assessment
- 3rd -4th week: Psychometric Features of Testing
- 5th -6th week: Assessing Language Skills

Despite the researcher's extensive efforts, only six participants completed all six weeks of the online workshop. Prior to the course, the participants were administered the Professional Knowledge of Assessment Test (PKAT) by Farhady (2016) to assess their current level of assessment literacy. To ensure alignment between the PKAT items, the Components of Theoretical Framework and the General Competencies for Teaching Profession developed by the National Ministry of Education, an expert was consulted.

Workshop:

The researcher meticulously prepared each session of the workshop using Prezi Video Program. To ensure the content was comprehensive and informative, the researcher reviewed numerous coursebooks in the field of language assessment and evaluation before developing the content for each week's session. Using Prezi Video Program, the researcher transferred the weekly content verbally and visually to create an engaging and interactive experience for the participants.

To ensure flexibility, the researcher emailed a video link of each session to the participants. This allowed them to attend the sessions at their own convenience, at any time that suited them. At the end of each session, the researcher provided a reference list and tasks designed to encourage participants to reflect on their current assessment practices and apply the new concepts they learned during the session.

Before the workshop sessions began, the PKAT was administered as a pre-test. This was done to establish the baseline assessment literacy levels of the participants and to track their progress throughout the workshop.

At the midway point of the workshop, the six participants were invited to a café to reflect on their progress. Unfortunately, due to their reluctance to undertake additional work in their classes, it was not possible to observe their in-class assessment practices. However, during the café meeting, the participants engaged in a semi-structured interview to provide feedback on the course, which helped the researcher gain valuable insights into how the participants were responding to the workshop content.

Throughout the course, participants were asked to write reflective notes after each lesson. These notes allowed participants to reflect on their learning and identify any areas of the content that they may have found challenging or needed further clarification. In response to participant requests, a revision video on the Psychometric Features of Testing was designed and shared with them. This video provided additional support to participants who needed extra help to understand the complex concepts covered in the workshop.

Upon completion of the workshop, the participants took the PKAT again as a post-test to track their progress and assess the impact of the workshop on their assessment literacy levels. This provided valuable data for the researcher to evaluate the effectiveness of the workshop and identify areas for improvement in future iterations.

Post Workshop

The primary objective of this phase is to evaluate the data collected using PKAT, reflective notes, and interviews. In the post-workshop phase, we aligned the items of PKAT with the components of the theoretical framework and the general competency indicators suggested by MoNE. We compared the results of PKAT scores with the alignment table to assess the impact of the course on participants' assessment literacy levels. Furthermore, we analyzed reflective notes and interviews to gain additional insights.

3.6. Data Analysis

Various analysis methods were used in the study. Quantitative data were subjected to statistical analysis, utilizing both descriptive and inferential statistics. On the other hand, qualitative data were analyzed using thematic analysis as the primary method. The following section describes the data analysis procedures applied in the study.

Firstly, the quantitative items suggested in PKAQ were aligned with the General Competency Indicators, and the content pages of the Assessment and Evaluation Course books were used to categorize the items under the main assessment-related topics. This alignment process was also reviewed by two experts. Next, a theoretical framework of language assessment literacy was developed for the study (see Appendix E).

The first part of the study aimed to understand the language assessment-related needs of participants. All volunteering English language teachers working at the secondary level in Mersin had filled out the PKAQ questionnaire. The questionnaire was analyzed through factor analysis to check its validity and reliability. Both exploratory and confirmatory factor analysis were used. After entering the quantitative data from PKAQ into SPSS, exploratory factor analysis was conducted to determine the number of factors that represented the items on the scale. The results were compared with the Theoretical Framework of Language Assessment Literacy components, and a match was not observed. Therefore, the theoretical framework was updated and revised. Another exploratory factor analysis was conducted to see item loadings and the number of factors. Following this step, a four-factor confirmatory factor analysis was conducted to check the construct validity of the questionnaire. Additionally, Cronbach's alpha was estimated for PKAQ. Since PKAQ has three quantitative items, each calculation and analysis was conducted independently three times.

The language assessment related needs of English Language Teachers were identified through the estimation of descriptive statistics including means and standard deviations, as well as percentages. These statistics were used to evaluate the importance, rate of knowledge, extent of classroom application, and other relevant factors related to language assessment literacy.

These findings were used to develop appropriate training programs and interventions to improve the language assessment literacy of the teachers.

To gain a deeper understanding of the needs of language teachers, a focus group interview was conducted through Zoom. The interview was transcribed and descriptive coding was used to categorize the data at a basic level. Braun and Clarke's six-step framework was then followed to analyze the data thematically. The transcribed data was read and reread to develop familiarity, and second initial codes were generated. Themes were then investigated around which codes could be organized. During the fourth step, the identified themes were reviewed to check their coherence. Once the themes were finalized, the results were written. Additionally, at the end of the Zoom meeting, the participants were sent the components of the theoretical framework of LAL for this study and asked to identify their needs by selecting from the topics given. The number of their replies on the components of language assessment literacy framework was also reported.

The second part of the study aimed to determine the effectiveness of the workshop language assessment course on the language assessment literacy levels of the participants. The pre- and post-tests were conducted using PKAT (Farhady, 2016), and the alignment of items between Competency Indicators and Theoretical Framework was checked before administering the tests (see Appendix F). The scores were analyzed using paired t-tests. Additionally, informal interviews were conducted at the end of the third week of the workshop. Reflective notes were written by the participants after each lesson, and the qualitative data gathered from these notes were treated as interview data and analyzed using Braun and Clarke's (2006) framework.

Overview of the data collection tools, the corresponding research questions, data analysis methods, can be seen from Table 4 below:

Table 4*Summary of research design*

Research Question	Data Collection Tools	Data Analysis
1. What are the language assessment literacy needs of EFL in-service teachers?	PKAQ Interview	Factor Analysis Descriptive Statistics Black and Clarke's Framework
2. What are the competency indicators that EFL in-service teachers have before the language assessment literacy (LAL) workshop?	PKAT General Competency Indicators	t-test Alignment of the PKAT Items with General Competency Indicators and Theoretical Framework
3. What competency indicators do EFL in-service teachers acquire through LAL workshop?		
4. Do LAL workshop effectively improve the language assessment literacy competencies of EFL in-service teachers?	Reflective Notes Semi-Structured Interview	Braun and Clarke's Framework

All in all, this study utilized both qualitative and quantitative data to answer the research questions.

3.7. Chapter Summary

This chapter provides an overview of the research project and the methods used for data collection and analysis. To ensure the reliability and validity of the data, both quantitative and qualitative data collection methods were employed.

Quantitative data was collected through the administration of the Pre-Test and Post-Test PKAT. The PKAT provided a numerical measure of the participants' assessment literacy levels before and after the workshop. The alignment table was used to assess the impact of the workshop on the participants' assessment literacy levels.

Qualitative data was collected through reflective notes, semi-structured interviews, and participant feedback. The reflective notes allowed participants to reflect on their learning and provided the researcher with insights into their experiences with the workshop content. The semi-structured interviews provided a deeper understanding of the participants' responses to the workshop content and allowed for the collection of more detailed information. Participant feedback was collected throughout the workshop to identify areas for improvement.

Both the quantitative and qualitative data were analysed using appropriate statistical techniques and qualitative analysis methods. The quantitative data was analysed using descriptive statistics to provide an overview of the participants' assessment literacy levels. Qualitative data was analysed using thematic analysis to identify common themes and patterns in the data.

The use of both quantitative and qualitative data collection methods allowed for a more comprehensive evaluation of the effectiveness of the workshop in improving the participants' assessment literacy levels. The reliability and validity of the data were ensured through the use of multiple data collection methods, providing a more complete picture of the participants' experiences with the workshop content.

4. RESULTS

This chapter presents the findings of a study that examined the language assessment needs of English Language Teachers working at the secondary level, as well as the impact of an online workshop on their language assessment literacy levels. The study employed multiple data collection methods, including the use of the Professional Knowledge of Assessment Questionnaire, the Professional Knowledge of Assessment Test, semi-structured interviews, and reflective notes. The Professional Knowledge of Assessment Questionnaire and Test were used to assess the teachers' knowledge and understanding of language assessment concepts, while the semi-structured interviews and reflective notes provided insight into their experiences with language assessment in their classroom practice.

4.1. Findings of the Professional Knowledge of Assessment Questionnaire

In this study, the language assessment needs of English language teachers at the secondary level were identified through the use of the Professional Knowledge of Assessment Questionnaire (PKAQ) and semi-structured interviews. To analyze the results of the PKAQ, factor analysis was conducted, and descriptive statistics such as means, standard deviations, and percentages were used to identify the teachers' language assessment related needs. Additionally, the results of the focus group interviews, along with the competence indicators and needs analysis, were considered in the design of the workshop syllabus. By taking a comprehensive approach to identifying the teachers' needs, the resulting workshop was better tailored to meet the specific needs of the teachers and help improve their language assessment literacy levels.

The responses to Item 9, which asked "What areas of language testing do you need to improve?" aligned with the findings of PKAQ and support its results. The theoretical framework that underlies the components of language assessment literacy, such as Basics of Testing (Methods of Test Preparation, Item Preparation), Assessing Language Skills (Assessing Speaking, Assessing Reading, Assessing Writing, Assessing Listening), Psychometric Features of a Test (Item Analysis, Knowledge of Statistics), and Functions of Assessment (Formative

and Summative Assessment, Placement Test, Portfolio Assessment) were identified as areas of improvement by the participants of PKAQ.

As seen in Table 5, in the evaluation of the psychometric properties of the PKAQ scale, which focuses on various dimensions of language assessment, Cronbach's Alpha was employed to determine the internal consistency of individual items. The scale demonstrated high reliability coefficients, with Item 10 ("the importance of language assessment related topics") yielding a Cronbach's Alpha of 0.948, Item 11 ("the knowledge and understanding of language assessment related topics") achieving a coefficient of 0.951, and Item 12 ("the extent of application of language assessment related knowledge") presenting a value of 0.945. These results indicate an excellent level of internal consistency for each item, suggesting that the PKAQ scale is a reliable instrument for assessing attitudes and competencies related to language assessment.

Table 5

Cronbach alpha values for PKAQ

Items of PKAQ	Value of Cronbach's Alpha
Item 10 the importance of language assessment related topics	0.948
Item 11 the knowledge and understanding of language assessment related topics	0.951
Item 12 the extent of application of language assessment related knowledge	0.945

A Principal Component Analysis (PCA) with Varimax rotation was used to explore the underlying dimensions of language assessment as captured by the PKAQ scale, focusing specifically on Items 10, 11 and 12 (see Appendix J for factor analysis). This methodological approach was selected to clarify the factor structure and to enhance the interpretability of the components.

For Item 10, the PCA extracted four distinct components, each representing a unique aspect of language assessment. Notably, 'Reliability of the Test' and 'Validity of the Test' emerged as highly significant factors, with loadings of 0.888 and 0.887 respectively. This finding underlines the importance of these constructs in the domain of language assessment. Other aspects such as 'Test Administration' and 'Assessing Grammar' also showed substantial loadings but on different components, indicating the multidimensional nature of the construct.

Similarly, the analysis of Item 11 revealed a multifaceted structure with several components extracted. The constructs 'Reliability of the Test' and 'Validity of the Test' again demonstrated high loadings, with values of 0.939 and 0.920, further confirming their key roles in language assessment. Other dimensions, including 'Assessing Reading Comprehension' and 'Deciding on What to Test', showed strong associations with different components, suggesting the presence of different factors within the language assessment field. These results indicated that language assessment, as conceptualized in the PKAQ scale, was not a unidimensional construct but rather comprises various interrelated dimensions. Each component identified in the analysis represented a specific aspect of language assessment, ranging from test design and administration to the evaluation of psychometric properties. This multidimensional framework offered an understanding of language assessment, which is crucial for both theoretical exploration and practical application in the field.

Likewise, the analysis of item 12 revealed distinct factor loadings across 27 items. In the initial matrix, key components such as 'Procedures in Language Test Design' (0.687), 'Deciding on What to Test' (0.710), and 'Writing Test Specifications/Blueprint' (0.673) predominantly loaded on the first factor. Notable loadings on other components included 'Assessing Grammar' (0.800) and 'Assessing Speaking' (0.870) on the second factor, highlighting the multifaceted nature of language assessment. Subsequent rotation further refined these relationships, with 'Writing Test Specifications/Blueprint' increasing to 0.841 on the first factor and 'Reliability of the Test' (0.939) and 'Validity of the Test' (0.920) showing strong loadings on the third factor. This analysis, converging in six iterations, underlined the complexity and distinct dimensions inherent in language test design and assessment.

Overall, the results of the factor analysis suggest that the PKAQ is measuring distinct constructs related to psychometric features of tests, basic testing concepts, assessing language skills, and functions of assessment. This information can be used to better understand the strengths and weaknesses of individual teachers' knowledge and to design more targeted professional development programs in the area of language assessment.

Table 6 below displays the responses of participants to three questions regarding their perception and application of language testing. The first column lists different topics in language testing, and the second to fourth columns show the number of respondents who rated the importance of the topic, their self-rated knowledge and understanding of the topic, and the extent to which they apply their knowledge in their class, respectively. The respondents rated the importance of each topic using a four-point scale: unimportant, not very important, fairly important and important. They also rated their knowledge and understanding of the topic and the extent to which they apply their knowledge using a four-point scale: very poor, poor, good, very good; rarely, sometimes, often, and always.

Table 6*Percentage distribution for PKAQ items 10, 11 and 12*

		PKAQ Item 10			PKAQ Item 11			PKAQ Item 12		
		Very Important	Fairly Important	Average	Very Good	Good	Average	Always	Often	Average
		%	%	%	%	%	%	%	%	%
1	Procedures in Language Test Design	44	48	46	20	65	42,5	21	44	32,5
2	Deciding on what to test	67	24	45,5	40	51	45,5	46	41	43,5
3	Writing test Specifications/Bl ueprints	40	54	47	29	51	40	27	42	34,5
4	Selecting Items for the Test	54	37	45,5	33	54	43,5	38	47	42,5
5	Preparing Items	51	44	47,5	34	52	43	36	45	40,5
6	Test Administration	47	42	44,5	30	49	39,5	34	37	35,5
7	Assessing Grammar	40	40	40	43	52	47,5	49	35	42
8	Assessing Pronunciation	42	42	42	24	43	33,5	28	35	31,5
9	Assessing Vocabulary	58	35	46,5	37	57	47	63	31	47
10	Assessing Reading Comprehension	62	31	46,5	43	48	45,5	58	31	44,5
11	Assessing Listening	63	24	43,5	30	46	38	35	38	36,5
12	Assessing Writing	59	28	43,5	37	47	42	38	41	39,5
13	Assessing Speaking	60	27	43,5	28	50	39	36	38	37

14	Formative Assessment	37	46	41,5	25	56	40,5	31	41	36
15	Summative Assessment	34	48	41	34	47	40,5	39	42	40,5
16	Diagnostic Assessment	35	50	42,5	30	51	40,5	26	39	32,5
17	Peer and Self-Assessment	43	41	42	31	48	39,5	27	43	35
18	Performance Assessment	46	44	45	34	49	41,5	38	36	37
19	Portfolio Assessment	45	36	40,5	28	52	40	25	38	31,5
20	Giving Feedback to Students	71	19	45	42	51	46,5	53	36	44,5
21	Developing Fair and Objective Scoring Rubric	68	25	46,5	31	53	42	40	42	41
Item Analysis										
(Item Facility (IF),										
Item Discrimination (ID),										
22	Choice Distractors (CD)	51	32	41,5	26	39	32,5	24	44	34
23	Reliability of The Test	62	24	43	27	53	40	38	40	39
24	Validity of The Test	63	22	42,5	30	49	39,5	34	44	39
25	Interpreting test scores	58	31	44,5	39	44	41,5	42	42	42
26	Knowledge of Statistics	40	43	41,5	25	46	35,5	25	40	32,5
27	Ethical Considerations in testing	56	29	42,5	33	49	41	38	48	43

Table 6 above provides percentages related to the language assessment needs of English language teachers. The high percentages for item 10 (“Which of the following topics in language testing is important enough to be included in a course on language testing?”) suggested that all components of language assessment literacy are important enough to be included in a language assessment course. However, when examining the table horizontally, there is a general trend of decreasing percentages indicating the knowledge and understanding level of the participants for each component. This emphasizes the need for a language assessment course that covers all of these components.

Compared to the average scores for item 10, the average scores for item 11 has a tendency to decrease that builds a strong base for low knowledge and understanding level of language assessment literacy of the participants. In line with the low average percentages of knowledge and understanding level of language assessment literacy of the participants, lower average percentages regarding the extent of the application supports the need for a training on language assessment literacy with all the language assessment literacy components of suggested.

To sum up, the tendency to decrease in the average percentages indicate that the participants believe in the significance of language assessment topics to be included in a language assessment related course however their knowledge and understanding prevent them from experiencing these components during their assessment procedure. This draws a vivid portrait for the language assessment related training need of the language teachers.

4.2. Findings on Language Assessment Needs

To gain a more comprehensive understanding of participants' language assessment related needs, a Zoom meeting was conducted on February 17, 2022. During the meeting, participants were asked to elaborate on their assessment-related needs as language teachers. During the focus group interview, participants were asked to share their language assessment related needs as language teachers. However, before delving into their assessment needs, participants expressed

their concerns about the problems they faced during their teaching which had an indirect impact on their assessment practices.

One common theme that emerged was the relationship between the curriculum, course books, and assessment. Participants highlighted the need for alignment among these three components in order to ensure effective teaching and learning. However, all of the participants expressed strong dissatisfaction with the course books they were required to use in their teaching.

Several participants shared their opinions on the course books, with some making the following comments:

The books are prepared by language teachers but they are very boring and complicated. Also, their content is too dense to be covered by both teachers and students. Unfortunately, we cannot think of assessment process apart from these course materials, and curriculum (Participant 2).

The course books are very boring and monotonous. We are at digital age and we need speaking activities. All skills should be focused (Participant 4).

Overall, participants felt that the course books they were using were not effective in supporting their assessment practices, as they did not provide adequate opportunities for students to develop their language skills and did not align with the learning objectives of the curriculum.

To further elaborate on their needs, participants also emphasized the impact of their target population's characteristics on their teaching and assessment practices. They all acknowledged that their professional knowledge level has decreased compared to their first year of teaching due to the language competency levels of their students. This has led to challenges in designing assessments that are appropriate and fair for all students. The participants discussed the need for additional training and support in designing and implementing assessments that meet the diverse needs of their students:

For three years, I started to question the reason why I am teaching English since English is the third language of most of my students. Due to the refugees from

Syria, most of the students are not competent enough to communicate in Turkish (Participant 1).

I want to confess that I am an English language teacher but I am not fluent in speaking English. I have not been abroad yet. I need training in Listening, Speaking and Pronunciation and I need to know more about assessing language skills, especially, listening and speaking (Participant 2).

Participants indicated that assessing skills is an essential component of assessment literacy that language teachers should possess. However, they highlighted that there are certain challenges that hinder their ability to develop their skills in this area. One of the challenges that they mentioned is the lack of proper equipment necessary to conduct effective skill assessments, which hampers their ability to accurately evaluate their students' language proficiency levels.

Another challenge that was mentioned is the low language competence level of the teachers themselves in the relevant skills. This makes it difficult for them to identify and evaluate the nuances and subtleties of the language skills they are assessing. Therefore, they emphasized the need for more training and professional development opportunities to improve their own language skills and become more competent in assessing language skills:

I am working as item developer at the center of Assessment and Evaluation Branch of Provincial Directorate of National Education in Mersin. They are asking me to write down an item to assess listening skill of the learner. I do ask them when I write down the listening task, isn't it a reading item ? This is a conflict, this is just to show the match between competency indicators of the curriculum and the test itself (Participant 3).

During the focus group interview, participants struggled to define assessment literacy in the context of their actual assessment practices. They found it challenging because the framework of assessment literacy did not align with what they were doing in their classes. The socio-cultural background of their students, their low language proficiency levels, and poor physical conditions for assessing skills added to their confusion. Therefore, the researcher modified the question and asked them to think about the components of assessment literacy that

they needed to know more about as English language teachers, regardless of their current teaching levels.

This approach yielded more focused and meaningful responses from the participants. They emphasized the need for further development in assessing language skills such as speaking, writing, and listening, as well as the use of technology in assessment. Participants also highlighted the importance of being able to design valid and reliable assessment tasks, analyze assessment results, and provide constructive feedback to students. They expressed their desire for training sessions and professional development opportunities to enhance their assessment literacy skills.

Upon analyzing the data, it was found that the most frequently mentioned component of language assessment literacy by the participants was Assessing Language Skills, followed by Function of Assessment and Basics of Testing, in descending order. However, Psychometric features of Assessment was reported as the least important component among the four components of the assessment literacy framework. During the focus group interview, some participants expressed their lack of use of psychometric features in their assessment practices, stating that they did not have sufficient knowledge and skills in this area.

One participant mentioned that they only use the basic statistics such as mean and standard deviation to analyze their test results, while another participant mentioned that they only have a superficial understanding of the concepts of reliability and validity. They also expressed the need for training and support to be able to use psychometric features more effectively in their assessment practices:

I have a table in which I try to check whether the students attain the related goal by putting + for the correct answer and – for wrong answers. Here, there is no numeric value (Participant 2).

Of course, I do consider the psychometric issues during my assessment practices, however I do not use numbers and make any kind of calculations (Participant 5).

It was evident from the discussion that there is a gap between the theoretical framework of assessment literacy and the actual assessment practices of the participants, and that there is a need for more training and support in this area.

After reviewing both the PKAQ results and the findings from the focus group interview, it became clear that language teachers in Turkey need training on various components of language assessment literacy. However, due to the limitations they face in terms of their students' proficiency levels, course books, and physical classroom conditions, they have limited opportunities to practice what they have learned.

It is worth noting that the curriculum for ELT in Turkish higher education does not provide any categorization regarding the level of students at primary education. This means that language teachers may not have a clear understanding of their students' language competency levels, making it more challenging for them to use the components of language assessment literacy effectively.

In light of these findings, a six-week language assessment literacy workshop has been designed to help language teachers in Mersin improve their knowledge and skills related to language assessment. This course aims to provide teachers with a better understanding of all the components of the language assessment literacy framework and equip them with practical tools and strategies to implement in their classrooms. By improving their language assessment literacy, language teachers can better support their students' language learning and ensure that assessments accurately reflect their students' language proficiency levels.

4.3. Findings on the Effect of Workshop on LA levels

The results of the PKAT were analyzed to determine the effect of workshop on the language assessment literacy levels of the participating teachers. In addition to these tests, informal interviews were conducted with the participants. At the end of each course, participants were asked to write reflective notes.

Table 7*Comparison of pre- and post-test results for LAL course*

	Pre-Test	Post-Test
Mean	13,17	16,16
Variance	54,56667	43,36
Pearson Correlation	0,780478	
df	5	
t Stat	-1,5667	
P(T<=t) two-tail	0,177966	
t Critical two tail	2,570582	

Table 7 presents the results of a pre-test and post-test for a group of participants in a language assessment literacy course. The mean score for the pre-test was 13.17 and the mean score for the post-test was 16.16, which indicates an increase in scores from pre-test to post-test. The variance for the pre-test was 54.56667 and the variance for the post-test was 43.36, which means that the scores for the post-test were less dispersed than the scores for the pre-test. The Pearson correlation coefficient between the pre-test and post-test scores is 0.780478, which indicates a strong positive correlation between the two sets of scores. The calculated t-statistic is -1.5667, which is less than the critical t-value of 2.570582. This means that the difference in means between the pre-test and post-test is not statistically significant at the 0.05 level. The p-value ($P(T \leq t)$ two-tail) is 0.177966, which is greater than 0.05. This also suggests that the difference in means between the pre-test and post-test is not statistically significant at the 0.05 level. In summary, the results suggest that there was an increase in scores from pre-test to post-test, but this difference was not statistically significant at the 0.05 level.

Upon completion of the first half of the workshop, the researcher organized a meeting with the participants at a cafe to discuss their experiences with the courses. During this meeting,

the participants shared their difficulties in following the workshop and explained the reasons behind their struggles to incorporate assessment practices in their classes. The following comments summarize their feedback:

Overall, the meeting provided an opportunity for the researcher to address these concerns and adjust the course content and delivery to better meet the participants' needs.

Participants who did not have an educational background reported difficulties in comprehending some of the assessment-related terms used in the workshop. This resulted in them having to watch the instructional videos multiple times in order to fully understand the concepts. These participants expressed the need for more simplified explanations of the terms used in the course materials to better support their learning. They suggested that including more visual aids, examples, and real-life scenarios could help them grasp the content more easily.

During the meeting with the participants, it was noted that all of them acknowledged the importance of the details presented in the workshop. However, they expressed that they faced difficulties in devoting sufficient time and attention to the courses due to their heavy administrative workload. As English Language Teachers, they had to fulfill various administrative duties such as grading papers, preparing lesson plans, and attending meetings, which made it challenging for them to fully engage with the workshop. Some participants suggested that the workshop courses could be made more flexible, allowing them to complete the assignments at their own pace, which would help them to balance their workload better. Others suggested that the administrative tasks should be reduced to provide more time for professional development, including the language assessment literacy course.

Participants expressed that the administrators in their schools were not fully aware of the significance of assessment in teaching and learning. This lack of awareness was seen as a hindrance to the implementation of effective assessment practices in the classroom. Teachers felt that there was a need for more support from the administration in terms of providing adequate resources and training to help them improve their assessment literacy.

Participants acknowledged that the use of rubrics during assessment was one of the most valuable points they learned through the workshop. They admitted that they were not using

rubrics while assessing speaking and writing skills, but the course raised their awareness of the importance of rubrics in providing clear assessment criteria and ensuring fairness and consistency in grading. They also stated that the rubric templates provided during the course were very helpful, and they planned to incorporate rubrics into their future assessments.

In their reflections, the participants also discussed the impact of the pandemic on their students' participation, particularly during speaking activities. They noted that despite their knowledge of effective assessment procedures, the current circumstances did not always allow them to implement the most appropriate methods in their classes. This highlights the challenge of adapting to the changing educational landscape and finding ways to continue delivering high-quality language instruction and assessment in a challenging environment.

During the meeting, the participants requested a revision of the psychometric features of the assessment, and one of the participants specifically asked for more information on Dynamic Assessment. They expressed a desire to deepen their knowledge in these areas, as they felt it would improve their assessment practices in the classroom. The course helped the participants to gain a better understanding of the importance of the instructions they used for assessment purposes.

In addition to the informal interview, the participants were also asked to complete a structured interview with three questions. They were asked to write down the things they liked and did not like about the courses, as well as three different things they had encountered during the courses.

Overall, the participants expressed satisfaction with the workshop and appreciated the opportunity to learn new things and refresh their knowledge of language assessment. They also expressed gratitude for being a part of the study. The psychometric features assessment tasks were noted as enjoyable and reflective by most participants. However, some participants had complaints about the length and level of detail of some courses, while others stated that they did not like the course on psychometric features of tests.

4.4. Findings on the Reflective Notes

To investigate the impact of workshop on language assessment literacy course, participants were instructed to reflect on their current assessment practices and what they have learned from the course. The reflective notes were guided by two main questions:

- What are the current assessment practices of English Language Teachers in their classes?
- What are the new assessment practices/methods that they learn from the online language assessment literacy workshop?

Based on the participants' feedback that they did not record the scores of their students, it was determined that asking about their current assessment practices related to psychometric features of testing would be irrelevant. Instead, they were provided with caricatures depicting significant features of psychometric principles violated by teachers and were asked to comment on them. This approach was intended to elicit their understanding and awareness of the importance of psychometric principles in language assessment, and to encourage them to reflect on their own assessment practices in light of this knowledge.

4.5. Findings on the Reflective Notes of Basics of Testing

At the beginning of the online language assessment literacy workshop, participants were asked to write reflective notes on the topic of designing tests. However, since they stated that they were not currently developing tests for their classes, they were instead asked to write a suggestion letter for a colleague, describing the steps involved in designing a test. All of the participants wrote letters that included the steps they had learned from the course, such as "The Purpose of the Test," "Use of Clear Instructions," "Preparing an Answer Key," and "Having Scoring Tables or Rubrics." Additionally, the importance of issues related to validity and reliability were emphasized by all participants.

4.6. Findings on the Reflective Notes of Functions of Assessment

In the second week, participants were provided a video on "Functions of Assessment" that described various types of assessments that could be used according to their purpose. At the end of the course, the participants were asked to write reflective notes focusing on their current assessment practices and the new things they had learned from the course. Five participants commented on their current assessment practices, and they emphasized that while they were aware of different assessment types, the low levels of proficiency of their students were a major factor limiting their use of different assessment types in their classes. The following comments were made by the participants regarding their current assessment practices:

As we are English teachers, I think we generally use different assessment types during our learning and teaching. For instance, I often use written exams, projects, course exams, individually questions and answers or game activities in my lessons, weekly working activities, word quizzes and like this. I also check of Ss' notebooks and give them personal feedback. (Participant 2, reflective notes on Functions of Assessment)

The 8th grade students will have an exam (LGS) at the end of 2nd term so they make repetitions as much as possible. That is why I use the quizzes. I randomly give the quiz papers to the students and we check them together, thus they can see their wrong learning. (Participant 4, reflective notes on Functions of Assessment)

It can be inferred from the comments made by the participants that they have a good understanding of their target population and are able to choose appropriate assessment types based on their students' needs and abilities. Some participants mentioned that they use formative assessments to monitor their students' progress and adjust their teaching accordingly. Others mentioned using summative assessments to evaluate their students' overall performance at the end of a unit or course. The participants also expressed their willingness to try new assessment types that they had learned about during the workshop, such as performance-based assessments and self-assessments. Overall, the participants demonstrated a good understanding of the

importance of selecting appropriate assessment types and the role of assessments in promoting student learning.

The participants in the study indicated that while they were already using a variety of assessment methods, the workshop provided them with new insights and approaches to assessment. Specifically, they mentioned that they were able to remember different assessment types better at the end of the course, and that they appreciated the emphasis on the importance of validity and reliability in assessments. It can be inferred that the course served as a refresher and reinforcement of their existing knowledge on assessment, while also introducing new ideas and concepts that they could apply in their teaching practice:

I know assessment types such as Self, Peer, Portfolio Assessment very well, I don't use them enough in my classes. Instead, I give more Project-based assignments. I realized that I could use alternative language assessments more and in a variety of ways. (Participant 1, reflective notes on Functions of Assessment)

I have learned new terms as types of assessments. In fact , I know the content but the terms are not familiar to me like CAT, CALT, CRT, NRT, KWHL, Dynamic Assessment. (Participant 2, reflective notes on Functions of Assessment)

After having listened your presentation, I realized that I wanted to measure not only my students' knowledge (traditional assessment) but also their language skills (alternative assessment). (Participant 4, reflective notes on Functions of Assessment)

At the end of the video I watched, I learned a new type if assessment: dynamic assessment. It is something we ignore as we teach. I also realized the variation of the assessment types I have used so far. (Participant 5, reflective notes on Functions of Assessment)

Based on the comments provided by the participants, it can be concluded that the workshop helped them to reflect on their current assessment practices and reminded them of the

different assessment types available. Although they acknowledged using a variety of assessment methods before taking the course, they emphasized that the course helped them to become more aware of the significance of assessment instructions, rubrics, and the psychometric features of assessment. However, since they did not want to take on additional work, it was not possible to observe how they applied what they learned in their actual teaching practices.

4.7. Findings of the Reflective Notes on Psychometrics Features of a Test

The participants' lack of interest in psychometric features of tests was evident, as they felt that they did not need to use any calculations for their assessment practices. Despite this reluctance, the workshop presented two parts on this topic. In part 1, the focus was on Reliability, Validity, Practicality, and Authenticity, while part 2 concentrated on Item Analysis, Scales of Measurement, Displaying Data, Describing Test Results, Normal Distribution, and Skewedness. Although the participants were asked to share their exams or assessment tools for a scenario during the workshop, they declined to share any documents. It is worth noting that although the participants did not use calculations, understanding the psychometric features of tests can help them improve the quality and effectiveness of their assessments.

To enhance the participants' understanding of the importance of psychometric features in assessment practices, the course shared five cartoons illustrating cases related to psychometric principles, and asked participants to provide feedback. Although the participants initially had some reluctance towards the topic, they acknowledged that they had encountered these scenarios in their own teaching experiences. Here are some of their comments:

Based on videos, cartoons and of course my previous knowledge, I am aware that terms such as reliability, validity, authenticity, practically etc. are essential parts for an assessment process. As I always say; being aware of them is not enough. We seriously need to put it into practice. (Participant 1, reflective notes on Psychometric Features of Test)

These cartoons are really funny and we encounter these situations very often. For example; in my first year of teaching, I prepared an exam that is unforgettable. I

prepared sections including all objectives of the units. The exam was too long, dense. The students couldn't finish on time. Preparing, administering and scoring the exam took lots of time .In short my test was not practical. It was lack of validity, reliability. Practicality. (Participant 2, reflective notes on Psychometric Features of Test)

Cartoon 1

Figure 8

Psychometric features of a test (Çeik, Çelik and Coombe, 2021, p. 44



The first cartoon demonstrated the issue of inconsistent grading and its impact on reliability. Participants acknowledged this as a common problem in assessment and some even admitted to having experienced it themselves. Here are some of the comments made by participants:

For example; as in the first cartoon, we ignore the reliability and evaluate according to the student. But I think I do this more in oral exams than in written exams. (Participant 1, reflective notes on Psychometric Features of Test)

In picture 1, there is a lack of reliability. But it can also be subjective. The situation seems unreliable from the point of students. The teacher can say that it is not true. Not to be subjective and to obtain more reliable results, I hide my

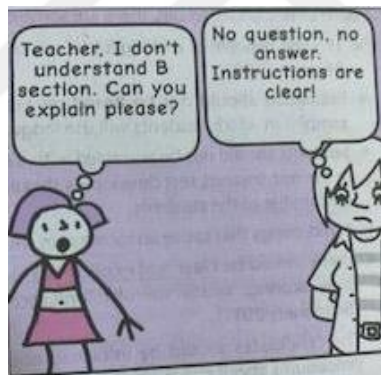
students' names while I'm evaluating the papers. (Participant 2, reflective notes on Psychometric Features of Test)

In the first caricature, students think that the teacher always gives high points to Rose. Maybe, this can be true or not; however, there is a mistake. That is because the teacher invalidates the reliability (Participant 4, reflective notes on Psychometric Features of Test)

Cartoon 2

Figure 9

Psychometric features of test (Çelik, Çelik and Coombe, 2021, p. 45)



In the second cartoon, participants were instructed to discuss the importance of providing clear instructions during assessment practices. They emphasized the negative effects that are lack of transparency, validity, and reliability can have on the assessment process. Interestingly, some participants shared that they had personally experienced similar situations to the one depicted in the cartoon:

The second cartoon is also very familiar to me, unfortunately I sometimes do that, too. Ok I won't do it again) (Participant 1, reflective notes on Psychometric Features of Test)

In cartoon 2, Instructions are not clear for students. There should be a transparency. It should provide clear and accurate information. I think there is

also a lack of test related reliability. These situations affect students' real performance. (Participant 2, reflective notes on Psychometric Features of Test)

The second caricature gives us some clues about validity. The validity is the second part of an exam because the reliability is prerequisite for validity. If an exam is validity, it has already the reliability. In the sample exam of this caricature, it can be understood the instructions are not clear so it is related to test-related reliability. Also it reflects there is a wrong performing in practicality. (Participant 4, reflective notes on Psychometric Features of Test)

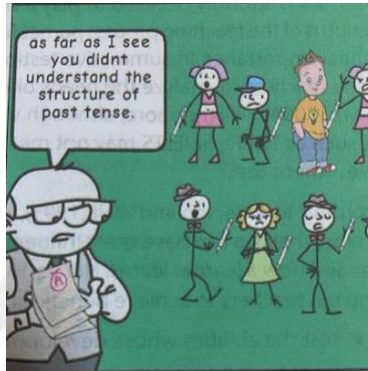
When I read the second visual, I have a smile. I sometimes do that. Because we, teachers are really get tired of explaining the instructions for many times. Especially students at lower grades don't listen while explaining the instructions. Then one asks what he doesn't understand, I explain it in a patient way, then the second but the same question comes, then the third ,.... Then I can lose my patience. I am really tired of telling the instructions. But it doesn't mean I am right about my feelings. We should be helpful to our students on every occasions. (Participant 3, reflective notes on Psychometric Features of Test)

Cartoon 3

In the third cartoon, participants were instructed to discuss the effects of test results on the motivation of the students.

Figure 10

Psychometric features of a test (Çelik, Çelik and Coombe, 2021, p.51)



The researcher intended to put emphasize on the backwash effect. Although all of the participants did not point the term backwash apparently, one way or another they stressed the effect of exam results on the students' motivation:

In cartoon 3, both teacher and students aren't satisfied with the results. We can say there is a negative backwash effect. The results affect the motivation of students and teacher. We can also talk about the validity of the test. (Participant 2, reflective notes on Psychometric Features of Test)

The teacher thinks that the students didn't understand the topic. Here it includes consequence-related evidence. There can be two possibilities: one of them is that students really couldn't learn anything about past tense. And this exam has the reliability, validity and practicality. The second one is that the teacher couldn't test students' learnings. This time this exam hasn't got any of them... (Participant 4, reflective notes on Psychometric Features of Test)

Cartoon 4

In the fourth cartoon, the participants discussed the emotional impact on students when their expectations are not met.

Figure 11

Psychometric features of a test (Çelik, Çelik and Coombe, 2021, p. 51)



The cartoon illustrates the negative backwash effect that this can have on students' motivation.

In cartoon 4, there is a lack of face validity. Teacher may have asked questions that students didn't expect. As a result, from students' perspective, it lacks face validity. So they say 'I will not study anymore.' The result of a test has a negative washback effect on students. They probably have low marks from exam. They didn't show their real performances although they studies hard. (Participant 2, reflective notes on Psychometric Features of Test)

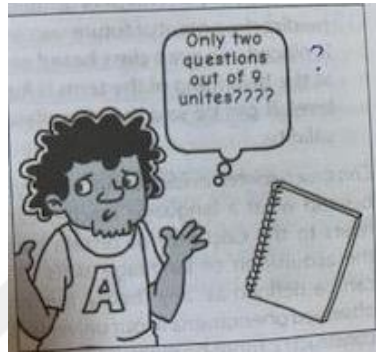
Learned helplessness is clearly seen in the 4th caricature. This makes us feel the reliability. In my opinion, learned helplessness damages the reliability. (Participant 4, reflective notes on Psychometric Features of Test)

Cartoon 5

The last cartoon is a good example of violation of content validity. The researcher aimed at having the participants pay necessary attention on the balance between the density of the content and the number of items on a test.

Figure 12

Psychometric features of a test (Çelik, Çelik abd Coombe, 2021, p. 47)



All of the participants could identify the violation and made extra explanation on the cartoon.

In the last cartoon there is a lack of validity. The test doesn't reflect a good representation of the content of the course. Intended criteria aren't measured correctly. Student complains about teacher's asking only 2 questions out of 9 units. (Participant 2, reflective notes on Psychometric Features of Test)

At first sight, content validity comes to my mind in this caricature. That's why there are nine units but only 2 questions... (Participant 4, reflective notes on Psychometric Features of Test)

This picture shows us the validity of the assessment, too. An assessment must be consisted all of the units which are taught to the students (Participant 6, reflective notes on Psychometric Features of Test)

The participants demonstrated a high level of analytical and reflective skills when analyzing and interpreting the cartoons provided to them. They were able to identify and explain the psychometric principles that were violated in each cartoon and acknowledged that they had experienced similar situations in their own assessment practices. Their comments revealed the significance of adhering to psychometric principles in both student learning and assessment practices. By following these principles, the results obtained from assessments can inform future

classroom practices and lead to better educational outcomes. The participants recognized that the consequences of not following these principles could result in negative effects such as a decrease in student motivation and inaccurate assessment results. Overall, the participants' insightful reflections emphasized the importance of psychometric principles in shaping effective and equitable classroom practices.

4.8. Findings on the Reflective Notes on Assessing Language Skills

In the 4th, 5th, and 6th weeks of the workshop, participants were introduced to the topic of assessing language skills, which was presented as a distinct module. During this module, participants were exposed to various methods and techniques for assessing language skills, and they were tasked with writing reflective notes for each skill.

In their reflective notes, participants were asked to draw on their existing assessment practices and note any new insights they had gained from the course materials. This exercise allowed participants to reflect on their own assessment practices and identify areas where they could improve their techniques for assessing language skills. By engaging in this reflective exercise, participants were able to better understand the importance of effectively assessing language skills in promoting student learning and achievement. They gained practical skills and knowledge that they could apply in their own teaching contexts, which would ultimately help to improve the quality and effectiveness of their assessments.

4.9. Findings on The Reflective Notes of Assessing Reading Skills

The participants discussed their approaches to assessing reading skills, taking into consideration the varying proficiency levels of their students. They noted that they used a range of assessment methods such as multiple-choice questions, short-answer questions, true/false statements, cloze tests, and reporting charts. When reflecting on their current assessment practices, the participants made the following comments:

While assessing reading, I usually use multiple choice, short answers and cloze tests. In the past it was easier to make them summarize a text or a short story. But it is more difficult now because we make them get used to give short answers or practice multiple choice. I also prefer and like true-false activities. Sometimes students who are very careful and focused are better than me and correct me and explain " No, teacher. It is true/false because it says.....". This really makes me happy. (Participant 3 reflective notes on Assessing Reading)

When it comes to reading, I must say that the students don't want read anything. They seldom look at the texts so they affect with panic and shy easily. I use close tests, multiple choices, and short answer mostly in exams, some mini quizzes. (Participant 4 reflective notes on Assessing Reading)

One of the participants stressed the point that she liked and would use in future while choosing a text.

In the video 'Points to consider' part is my favorite. It raised my awareness. I will be more careful with the texts I 'm using. (Participant 2 reflective notes on Assessing Reading)

The participants noted that they use various assessment methods such as cloze tests, multiple-choice questions, short-answer questions, and true/false statements to assess reading skills. However, they also highlighted that due to the low proficiency levels of their students, they are limited in their ability to use more diverse assessment methods. In addition, the participants did not provide any further comments on their current assessment practices, aside from noting the new insights they gained from the workshop. This lack of commentary suggests that the participants may not have critically examined their current assessment practices or considered alternative approaches to assessing reading skills.

4.10. Findings on The Reflective Notes of Assessing Speaking Skills

Based on previous data gathered through focus group interviews, speaking was identified as the least practiced skill among the four language skills. This was attributed to various factors such as students' low proficiency levels in both Turkish and English, as well as the diverse sociocultural backgrounds of the participants that may limit or hinder them from practicing speaking skills with their students.

Despite these limitations, the participants listed various examples of speaking assessment tasks that they use in their teaching, including controlled and guided dialogue completion methods, question and answer activities, imitation exercises, role plays, discussions, and group conversations. These assessment tasks aim to assess students' speaking skills and provide opportunities for them to practice their speaking abilities in a supportive learning environment:

I think speaking is the least practiced skill in schools in our country. Language teaching is generally limited to writing and reading. In my classes, I have been using controlled and guided dialogue completion in accordance with the textbook and curriculum. Question and Answer method is one of the methods I use most in the classroom. (Participant 1 reflective notes on Assessing Speaking)

Students generally have difficulty in a foreign language especially in speaking. That is why they feel embarrassed and as if everybody will laugh at them while talking etc. Therefore, it is a bit hard for us to force them speaking. Sometimes, we practice it with plays, games or activities. Also, there are alternative speaking activities in course books. We use them actively. (Participant 4 reflective notes on Assessing Speaking)

During the workshop, it was noted that some participants did not use rubrics to assess their students' speaking skills during presentations, discussions, and other activities. One participant even admitted to not using any type of rubric for this purpose. However, other participants did not comment on their use of rubrics for assessing speaking skills. A new topic that was raised

during the course was the use of scoring scales in language assessment. This topic may have provided participants with valuable insights into effective assessment practices and helped them to better evaluate their own assessment methods. Following comments were made:

There are alternative speaking activities in course books. We use them actively. However, that is not enough. When I think on my own, I feel secure; I use classroom language mostly with gestures or mimes, let students talk to each other via dialogues, games. Firstly, they try to understand what I am talking about, then keep the meaning in time and actively participate in the progress. They specially imitate my pronouncing... I generally evaluate them verbally and do not use any kind of rubric or checklist. (Participant 4 reflective notes on Assessing Speaking)

In this video, SHOULD and SHOULDN'T parts and scoring Scales are very useful for me. After that, I decided to prepare scoring scales beforehand and measure their skills in a right way. (Participant 2 reflective notes on Assessing Speaking)

During the workshop, participants emphasized the various limitations that hindered them from using different methods to assess their students' speaking skills. These limitations included classroom size, students' readiness and proficiency levels, and their diverse sociocultural backgrounds. Despite these limitations, the participants indicated that they do assess their students' speaking skills, albeit through a limited range of methods.

One new concept that was emphasized during the course was the use of scoring rubrics in language assessment. Participants learned about the benefits of using rubrics to provide clear criteria for assessing speaking skills, and how this can help to ensure fairness, accuracy, and consistency in grading. By incorporating rubrics into their assessment practices, participants may be able to overcome some of the limitations they face in assessing speaking skills and provide their students with more effective and targeted feedback.

4.11. Findings on The Reflective Notes of Assessing Writing Skills

During this part of the workshop, participants were introduced to different methods for assessing writing skills. Unlike speaking, writing was reported to be more commonly practiced in classes, with participants noting that students often feel more comfortable during writing exercises. However, the level of competency of the students remained a major factor that influenced the participants' choice of assessment practices for writing.

Interestingly, none of the participants reported using scoring rubrics while assessing their students' writing skills. This echoed the findings of the reflective notes on assessing speaking, which also noted a lack of rubrics being used. However, the significance of using rubrics when assessing writing was emphasized as a new concept learned from the workshop, and all participants expressed their gratitude for the new writing tasks that were introduced during the course:

Our students write letters, invitation cards, dialogues concerning objectives of the unit, summaries of short stories that they read, prepare poster, brochures. They are in imitative level. They write and I correct the mistakes. They learn the right ones. They make so much mistakes like punctuation, spelling, and grammar. I support them until they write the right ones. As you said, we use nearly all of the sample tasks you mentioned in the class .While we are doing and choosing activities, I give attention to my students' competency level and keep in mind my purpose of writing. We make controlled guided writing. We use keyword writing ,picture writing ,reordering most. Because they are suitable to my students' level. (participant 2 reflective notes on Assessing Writing)

Students feel more confident than the other skills at this part. In fact they like free-writing. I always tell them to change the language as English on computers, on mobile phones...They write something for English Corner related to the topic at the end of the unit. (participant 4 reflective notes on Assessing Writing)

During the course, a participant shared her approach to scoring writing assessments, highlighting the importance of spelling, grammar, vocabulary, and meaning. She further mentioned that she had not utilized any type of scale or rubric in her assessments. Consequently, the participant regarded the use of scoring scales as a new concept that she learned from the workshop:

In fact, while watching your video, I have realized that I haven't scored my students' writing tasks. I have been giving writing tasks to my students... They read what they write then I try to correct the mistakes but haven't marked or scored them... (Participant 3 reflective notes on Assessing Writing)

In my classes, as I'm a teacher of fifth and sixth grades we have very limited writing tasks. In fifth grades, I generally make my students do imitative and intensive level writing tasks. I score them according to spelling, grammar and meaning. (Participant 3 reflective notes on Assessing Writing)

I liked the writing scale in this session, I'd be happy to use that in my classes in the future. I also learned new kinds of tasks. Thank you. (Participant 5 reflective notes on Assessing Writing)

Based on the shared reflections, it is apparent that teachers vary their writing assessments based on the level of their students. Nonetheless, none of the participants mentioned the use of rubrics in their current assessment practices. Through the workshop, the use of scoring rubrics was introduced as a new approach to writing assessment, and all participants expressed appreciation for this novel concept. They even showed willingness to incorporate rubrics in their future assessments.

In conclusion, the workshop on assessing language skills provided valuable insights and new techniques for teachers to improve their assessment practices. Participants realized that they had been neglecting the importance of scoring their students' writing tasks and were grateful for the introduction of a writing scale in the course. Despite the limitations and challenges they faced, such as the low proficiency levels of their students, participants showed enthusiasm in implementing new types of writing tasks and incorporating rubrics in their assessments. The

course served as a platform for participants to reflect on their current practices and learn new techniques that will benefit their students in the future.

4.12. Findings on The Reflective Notes of Assessing Listening Skills

The final component of the workshop on assessing language skills focused on assessing listening. It was noted by the participants that listening is often a neglected skill due to factors such as crowded classrooms and poor physical conditions. As a result, it is not assessed thoroughly as compared to other language skills. Following comments regarding assessment practices of the participants were made:

In school books there are lots of listening activities and this is perfect. I generally do it in my classrooms. However, we don't use this skill in any kind of exam. We can order a lot of reasons: ... so many contents in each unit, very crowded classrooms, volume and the time are not enough for hearing. Thus, they always want additional time and the exam can't be finished on determined time. They can't take notes while listening the text or dialogues. They want to listen again and again. They fail at the end. (Participant 4 reflective notes on Assessing Listening)

The biggest challenge we have is the unmotivated students. Most of the time we, teachers, have to spend more time to get their attention on the task. (Participant 5 reflective notes on Assessing Listening)

Despite the challenges posed by crowded classrooms and poor physical conditions, some of the participants who have access to smart boards and smaller class sizes have found ways to incorporate a variety of assessment tasks for listening. They emphasized the importance of assessing listening skills, as it is a crucial part of language learning and communication. They mentioned using different types of listening tasks such as multiple choice questions, note-taking, gap-filling, and summarizing to evaluate their students' listening abilities. However, other participants who faced difficulties in creating a suitable listening environment in their classes

expressed their interest in learning new methods to assess listening skills, as they acknowledged the significance of this skill for their students' language development.

I must say that the listening activities are limited to the listening tracks in the textbook and the extra-curricular activities that we do on the interactive whiteboards. I can say that I usually use Intensive and Selective Listening activities in secondary school. Responsive and Extensive Listening includes higher level activities, but I also use them in my lessons. Dictation is my favorite☺ but I don't use any detailed assessment for listening. (Participant 1 reflective notes on Assessing Listening)

Thanks to smart boards, our listening parts are great anymore. In units' presentation, listening parts also include subtitles or written text... I use selective listening most, before listening, I make them read the questions about the text. Then I make them listen. As the students have read the questions beforehand, they search for specific information. I can use open ended response to a question here. (Participant 2 reflective notes on Assessing Listening)

In my classes I use both intensive and extensive listening tasks. We have resources and tools so we can do the listening tasks in the language classes. Our classes consist of 25 students max so we can say the classrooms are not that crowded for listening. (Participant 5 reflective notes on Assessing Listening)

The participants highlighted that listening is one of the skills that is least practiced and assessed in language classrooms due to various reasons such as crowded classrooms, poor physical conditions, and limited resources. However, some participants with access to smart boards and smaller class sizes reported using a variety of assessment tasks for listening. Despite this, only one participant mentioned learning about authentic tasks through the workshop, which could potentially enhance the assessment of listening skills in their classrooms.

4.13. Findings on The Reflective Evaluation of the Workshop

The participants were requested to share their thoughts on the workshop, and most of them mentioned that the videos were informative and presented in a clear manner. Some participants also mentioned that the videos acted as a refresher of their prior knowledge about assessment. Following comments were made:

Thank you for all your presentations. I want to talk about all videos. They are useful ones. Giving sample tasks have made it clearer and funnier. I learnt many details about assessing four skills. But to tell the truth, our students' level of proficiency are very low. As a result, measuring students' four skills can't be made correctly. Now, let's look at how I measure my students' four skills. (Participant 2)

You enabled me to remember the assessments. In fact while watching the videos I can admit that I have forgotten the terms about assessments and you helped me to think that how I can be a better teacher or how I can make my students be a better student. I have been thinking about the ways you have I used in the past but not using anymore. The process was very educative and informative for me. Thank you for your great effort. (Participant 3)

Participants were asked to reflect on their experience with the workshop on assessment, and many of them expressed that the videos were informative and presented the content clearly. They also mentioned that the courses served as a helpful reminder of previous learning related to assessment. However, beyond simply reminding them of what they already knew, the courses also provided participants with new information and allowed them to delve deeper into the function of assessment. As a result, they felt more confident in their ability to assess their students' skills and believed that they would be able to teach their students better.

I would like to thank you for this training. Actually, I feel bad that I haven't known about most features of testing until now. We had testing course back

in college but I don't remember. The only thing I remember is the variety of tasks. (Participant 4, reflective notes on Functions of Assessment)

I believe most teachers share the same opinions. My point of view about the reason of this is that our teachers did not tell about the importance of testing. We just saw that as a formality to run. After this training I feel more aware and ready for testing. I have learned about different types of testing, psychometric features and the task types. The most important part is, I think, we thought about the objective of testing. (Participant 5)

To sum up, assesment is the indispensable chain of learning and teaching process that has profound impact on one's life. As a result teachers' keeping up the recent changes in their field means a lot. Providing necessary support through workshops, inservice training sessions is vital for in-service teachers.

4.14. Effect of Online Workshop on Development of Competency Indicators

Two of the research questions in this study focused on the impact of an online language assessment literacy workshop on the competency indicators proposed by MoNE. As the competency indicators do not provide numerical values for competency levels, the researcher examined the alignment between the components of the theoretical framework, the competency indicators, and the items of PKAT. To ensure accuracy and reliability, an expert from the Branch of Assessment and Evaluation Department was consulted to review the alignment independently.

After listing the pre and post-test scores of the participants, the correct and incorrect answers for each test were counted separately. The total number of correct answers for each component of language assessment literacy was calculated by considering both pre and post-test answers of the participants.

Table 8

Alignment between pre-test and post-test results of online language assessment workshop

Components of Theoretical Framework	Competency Indicators	Total Number of Correct Answer-Pre-test	Total number of Answers for Pre-Test	Total Number of Correct Answer-Post-test	Total Number of Answers for Post Test
Functions of Assessment	B4.1/B4.2	30	78	42	78
Basics of Testing	B4.1/B4.2/B4.3 /B4.4/B4.5	23	42	18	42
Psychometrics	B4.3/B4.4	12	54	13	54
Assessing Language Skills	B4.1/B4.2/B4.3	14	36	16	36

Table 8 shows the results of the pre-test and post-test for the different components of the theoretical framework and competency indicators. The components are Functions of Assessment, Basics of Testing, Psychometrics, and Assessing Language Skills. The competency indicators for each component are listed next to it. The table shows the total number of correct answers for each component in the pre-test and post-test, as well as the total number of answers for each test. For example, for the Functions of Assessment component, there were 30 correct answers out of 78 total answers in the pre-test, and 42 correct answers out of 78 total answers in the post-test.

Overall, the post-test results show an improvement in the number of correct answers for all components, with the largest improvement seen in the Functions of Assessment component. The pre-test results suggest that participants had some knowledge of the different components,

but the post-test results indicate that the online language assessment literacy workshop helped to improve their understanding and competency in these areas.

Despite this limitation, it is worth noting that the total number of correct answers for the pre-test and post-test suggests that the workshop had a positive effect on the development of the competency indicators. However, it is important to acknowledge that the number of participants in the study was limited, and therefore, the results cannot be generalized beyond this group of participants. It is also worth noting that the competency indicators were not assessed on a level basis, which further limits the ability to draw conclusions about the effectiveness of the course.

4.15. Chapter Summary

The results of the study suggest that the participating teachers had varying levels of understanding and knowledge regarding language assessment concepts, and that they had a clear need for professional development in this area. The workshop was found to have a positive impact on their language assessment literacy levels, with the teachers reporting increased confidence and competence in designing and administering assessments, as well as analyzing and interpreting assessment data.

Overall, this study highlights the importance of providing ongoing professional development opportunities for English Language Teachers working at the secondary level, particularly in the area of language assessment. The use of multiple data collection methods provided a comprehensive understanding of the teachers' needs and the impact of the workshop on their professional development.

5. DISCUSSION AND CONCLUSION

The aim of this study was to identify the language assessment-related needs of English Language Teachers working at the secondary level in Mersin and develop an online Language Assessment Literacy workshop to improve their language assessment skills. The study comprised two parts: needs analysis and course design, and the impact of the workshop. The needs analysis was conducted using the Professional Knowledge of Assessment Questionnaire and focus group interviews to determine the components of the language assessment literacy framework to be covered in the course. These components were aligned with the general competency indicators of MoNe and the Professional Knowledge of Assessment Test, which were used as pre- and post-tests for the course.

The second part of the study aimed to assess the effectiveness of the online language assessment workshop. To evaluate the impact of the course, the Professional Knowledge of Assessment Test, interviews, and reflective notes were used. However, due to ethical considerations, the teachers did not provide samples of their in-class practices. Therefore, the study was unable to observe changes in the teachers' assessment literacy levels before and after the course. Additionally, the teachers did not wish to engage in any extra in-class work for the study.

The limited number of participants in the study means that the data obtained from the pre- and post- PKATs is insufficient to explain the effect of the course on changes in the participants' language assessment literacy levels. Nevertheless, the findings from interviews and reflective notes suggest that the language assessment literacy course increased participants' awareness of assessment-related topics, such as the steps involved in test design, various assessment methods, the purpose of assessment, and psychometric features of tests, especially reliability and validity issues. In particular, the workshop was highly beneficial for participants in terms of using scoring rubrics during language assessment procedures.

Although the number of participants was small, the alignment table between PKAT and General Competency Indicators suggests a positive change in the language assessment literacy

levels of participants. However, it is important to note that the results cannot be generalized due to the limited sample size.

The remainder of this chapter will briefly discuss significant issues raised at the end of the study. After that, the pedagogical implications are explained. The final section will include limitations of the present study and suggestions for further research.

5.1. Analysis of Workshop Efficacy

The substantial improvement in language assessment literacy among in-service EFL teachers post-workshop warrants a deeper analysis. This improvement signifies more than just an increase in theoretical understanding; it represents a practical enhancement in the teachers' ability to assess language proficiency effectively. Such an outcome is particularly significant in light of studies reporting the impact of professional development programs on teachers' language assessment literacy (Baker & Riches, 2017; Boyd & Donnarumma, 2018; Kremmel et al., 2018 cited in Giraldo, 2021). The study's results echo these findings but add a new dimension by demonstrating the effectiveness of an online medium, a relatively unexplored territory in teacher education.

Furthermore, the data reflects not just a marginal improvement but a marked transformation in some cases, suggesting that the workshop content was not only well-received but also effectively internalized. The nuanced improvement across different assessment literacy components also indicates that the workshop successfully addressed diverse learning needs. This aligns with Jones and Blankenship's (2017) study, which highlights the importance of differentiated instruction in adult learning. The workshop's design, which likely included elements of interactive learning, practical exercises, and, appears to have been pivotal in achieving these outcomes.

In conclusion, these findings underscore the significant potential of online professional development to enhance language assessment literacy. They contribute to a growing body of research advocating for the integration of digital technologies in teacher education, suggesting

that with carefully designed content and delivery methods, online platforms can be as effective, if not more, than traditional face-to-face training programs.

5.2. In-Depth Exploration of Diversity in Learning Progression

The study's findings on the diversity in learning progression among teachers offer profound insights into the complexities of professional development. This diversity reflects the individualized nature of learning, as each teacher brings unique experiences, prior knowledge, and personal teaching contexts to the learning process. The variability in progression aligns with Xu and Liu's (2009) observations on individual learning trajectories in educational interventions. This suggests a departure from traditional, uniform approaches to professional development, emphasizing the need for more personalized, adaptive learning experiences.

Moreover, the study's findings reiterate the significance of understanding and catering to different learning styles in professional development programs. As highlighted in Johnson's (2016) research, adult learners benefit from approaches that recognize and adapt to their unique learning preferences and needs. This is particularly relevant in the context of language education, where teachers often face diverse challenges and require specific skill sets to address them effectively.

The implications of these findings are far-reaching, underscoring the importance of flexibility and customization in professional development for teachers. They advocate for a shift towards more learner-centered training models, which can accommodate the varied pace and style of adult learning, thereby enhancing the overall efficacy of such programs.

In light of these insights, future professional development initiatives should consider incorporating diagnostic tools to assess individual learning needs and preferences. This can lead to more targeted and effective learning experiences, ultimately contributing to more competent and confident teaching practices in language assessment.

5.3. Comprehensive View on Assessment Literacy Needs

The initial assessment phase of the study, focusing on the diverse literacy needs of teachers, is a key aspect of its findings. This phase revealed a broad spectrum of pre-existing competencies and gaps in understanding language assessment. This observation is particularly relevant when considered alongside Inbar-Lourie's (2013) advocacy for a needs-based approach in language teacher education. By identifying these varying needs, the study was able to tailor the workshop content more effectively, ensuring relevance and applicability to the participants' real-world teaching scenarios.

The diversity in literacy needs also echoes broader educational research, which emphasizes the importance of personalized learning paths in professional development. Understanding teacher's need is one of the crucial steps for designing effective training programs (Roberts, 1998). This tailored approach not only ensures that the program addresses specific gaps in knowledge and skills but also fosters a more engaging and relevant learning experience for teachers.

Furthermore, the study's approach aligns with contemporary educational paradigms that prioritize learner-centered methodologies. This is particularly important in the context of adult education, where learners come with varied professional experiences and expectations. The success of the workshop in enhancing assessment literacy among participants underscores the effectiveness of this personalized approach.

In sum, these findings support a paradigm shift in professional development for educators, advocating for the adoption of more individualized and needs-based training models. This approach has the potential to significantly enhance the efficacy of professional development programs, particularly in fields that require specialized skills such as language assessment.

5.4. Detailed Exploration of Workshop Structure and Content

The structure and content of the online workshop, as evidenced by the study's findings, played a pivotal role in the enhancement of language assessment literacy among participants.

This is consistent with Taylor's (2009) assertion that effective professional development should marry theoretical knowledge with practical application. The workshop's content was not merely theoretical but included practical, real-world applications, crucial for adult learners.

The interactive elements of the data collection tools, such as interviews, likely contributed significantly to the learners' engagement and understanding. This is in line with Vygotsky's (1978) social learning theory, which emphasizes the importance of collaborative learning. By integrating these interactive components, the workshop likely facilitated deeper learning and understanding among participants.

Moreover, the workshop's focus on real-life scenarios and practical application of language assessment techniques is crucial. It reflects the research by Johnson and Johnson (1999), which highlights the effectiveness of experiential learning in adult education. By engaging teachers in activities that closely mirror their professional challenges, the workshop ensured that the learning was not only theoretical but also immediately applicable in their teaching contexts.

In essence, the structure and content of the workshop were instrumental in achieving the observed improvements in language assessment literacy. This underscores the importance of thoughtful curriculum design in professional development programs, particularly in the context of language education, where practical application is as important as theoretical knowledge.

In conclusion, this study's comprehensive analysis of an online language assessment literacy workshop for in-service EFL teachers has provided crucial insights. The effectiveness of the workshop format in enhancing assessment literacy, the diversity in individual learning progressions, the importance of addressing specific literacy needs, and the impact of a well-structured, content-rich program have all been highlighted. These findings collectively underscore the potential of online professional development platforms and point towards a more personalized, practical, and interactive approach in teacher education, particularly in the domain of language assessment.

5.5. Pedagogical Implications

This present study has several important implications for in-service training, teacher education programs, and language education in Turkey.

The findings of the study highlight the need for more comprehensive in-service training programs for English Language Teachers (ELTs) in Turkey. The results indicate that EFL teachers lack sufficient language assessment literacy, which is crucial for effective teaching and learning. Therefore, it is recommended that relevant stakeholders, such as the Ministry of National Education (MoNE), should prioritize the development and implementation of training programs that focus on enhancing their assessment literacy skills. The challenges underscore the need for in-service training programs to be designed based on a careful examination of the needs and characteristics of the target population. It is important to take into consideration teachers' past experiences and concerns when designing training programs to ensure that they are relevant and meaningful. Additionally, efforts should be made to minimize the burden on teachers by considering their workload and time constraints. By taking a more strategic and tailored approach to in-service training, it is more likely that teachers will be willing and motivated to participate, and ultimately, benefit from such programs.

The study's implications extend to teacher education programs in Turkey. The results indicate that there is a need to incorporate more explicit and comprehensive assessment literacy training within teacher education programs. This will help to ensure that future teachers possess the necessary skills and knowledge to effectively assess students' language skills and promote successful language learning outcomes.

The implications of the study relate to language education in Turkey more broadly. The results suggest that there is a need for a shift towards a more balanced approach to language education, where assessment is integrated into the curriculum and is not viewed solely as a separate component. This approach would help to ensure that language assessment is a meaningful and integral part of language education in Turkey.

The study has important implications for addressing language teachers' challenges in their teaching and assessment practices. In a meeting held by the Provincial Directorate of National Education in Mersin, several English Language teachers expressed their dissatisfaction with the course books distributed by MoNe, stating that they are dull and fail to capture students' interest. In addition, teachers teaching in classes with students of lower Turkish proficiency faced communication difficulties, which hindered their ability to deliver effective instruction and led to a loss of motivation and questioning of their role in the classroom.

These issues highlight the problem of burnout and reduced motivation among language teachers to engage in professional development and enhance their teaching and assessment practices. To address these challenges, the components of language education in Turkey must be examined in greater detail, taking into account students' diverse socio-cultural backgrounds, materials, and classroom equipment. More specifically, the quality and effectiveness of course materials should be reviewed, and strategies for addressing communication difficulties between teachers and students with lower Turkish proficiency should be developed. By addressing these issues, language teachers can feel more supported and motivated, which can ultimately lead to more effective teaching and learning outcomes for students.

The present study also sheds light on the ambiguous role of assessment in language education. While some teachers view assessment as an integral part of their teaching practices and use it effectively in every phase, others simply consider it as a routine procedure to be followed without much attention or care. This disparity highlights the need for more comprehensive and systematic training on language assessment for teachers, especially in terms of developing their assessment literacy. Moreover, there is a need to enhance the collaboration and communication between language teachers, assessment experts, and educational researchers to establish a more unified and coherent approach to language assessment in Turkey. By doing so, the quality and effectiveness of language education can be improved, and students' learning outcomes can be maximized.

In addition, it was observed that assessment is often viewed as a routine procedure that teachers must follow by both teachers and administrators. However, assessment plays a crucial role in various aspects of language education, beyond just marking papers. During an informal

conversation with the head of the Assessment and Evaluation Branch at the Provincial Directorate of National Education in Mersin, she expressed concern about the lack of attention given to assessment by managers. This highlights the need for a more comprehensive understanding of assessment, and the importance of promoting assessment literacy among both teachers and administrators.

The perception of assessment by all stakeholders is another important implication that emerged from this study. The interviews conducted as part of this research highlighted the need to address this issue. It is essential to explore and understand how different stakeholders, including teachers, students, parents, and administrators, perceive the role and purpose of assessment in language education. These perceptions can significantly impact the effectiveness of assessment practices and their outcomes. Therefore, it is crucial to involve all stakeholders in discussions and decision-making processes related to assessment practices in language education.

The study sheds light on the limited coverage of assessment-related courses in Education Faculties and English Language Teacher Education programs. Typically, these programs only offer two courses on assessment and evaluation, and some related courses start as late as the 8th semester. Additionally, there is only one elective course that covers in-class assessment practices. This implies that there is a need to revisit and expand the current assessment curriculum in teacher education programs to equip future teachers with a stronger foundation in assessment literacy."

The significance of assessment should not be limited to the responsibility of the teacher, but should be recognized by all stakeholders in the field of education. It is crucial to emphasize the importance and interrelation of teaching and assessment from the early stages of teacher education across all fields. By doing so, the perception of assessment by practitioners can change, and classroom practices, which play a vital role in one's life, can improve accordingly.

Furthermore, this study has significant implications for the general competency indicators proposed by the Ministry of National Education (MoNE). One key issue is that the section on teacher competency indicators related to assessment lacks clear criteria and scoring methods, making it difficult to objectively assess teachers' performance (MONE, 2017). This is a concern,

as assessment is a critical component of effective teaching and learning, and teachers need to be equipped with the necessary competencies to design, implement and evaluate assessment tasks that align with curriculum objectives and promote student learning (Black & Wiliam, 1998; Popham, 2009). Thus, it is essential to provide teachers with clear and measurable competency indicators related to assessment that align with best practices in the field.

A crucial implication of this study concerns the professional development of language teachers. Initially, the study involved 10 volunteers, but as they learned that the courses would be conducted in English, some of them felt that they may not possess the necessary language proficiency and subsequently withdrew from the study. A few of the teachers expressed that they would have liked to continue but did not feel confident enough. This highlights a significant issue that language teachers face - a lack of motivation to develop their professional skills, particularly in language competency. The teachers' hesitance to share their assessment practices can also be attributed to their perceived inadequate English proficiency.

In summary, the study's findings have important implications for improving in-service training, teacher education programs, and language education in Turkey. By prioritizing the development and implementation of comprehensive assessment literacy training programs, stakeholders can ensure that teachers possess the necessary skills and knowledge to effectively assess students' language skills and promote successful language learning outcomes.

5.6. Limitations

Although the results of PKAQ and focus group interview built a strong basis for the identification of the language assessment related needs of language teachers and the design of the workshop, the results of PKAT, interviews and reflective notes cannot be generalizable due to the limited number of volunteer teachers taking part in the study.

After obtaining necessary permissions from the provincial directorate of National Education in Mersin, the researcher encountered difficulties in recruiting participants for the study. Despite attending several meetings held on different dates to present the study and seek volunteers, teachers were hesitant to even fill out the questionnaires. Some expressed frustration

about the frequency of requests for participation in graduate studies, and became even more reluctant when informed that they would be asked to participate in in-service training. In fact, one teacher remarked that they were tired of attending in-service trainings that had little to no impact on their teaching practices.

Due to the voluntary nature of the study, the language teachers did not exhibit any noticeable change in their assessment practices. Additionally, the participants were not willing to share their existing practices or engage in any extra work during the workshop. It was not possible to determine the impact of the workshop on the different levels of language proficiency, as the General Competency Indicators did not have a defined scoring table with associated levels. Therefore, it was only possible to demonstrate the change in scores between the pre- and post-tests.

Despite obtaining the necessary permissions from the Provincial Directorate of National Education Mersin, the majority of language teachers expressed disinterest in participating in the study. They cited dissatisfaction with previous in-service training programs, reluctance to fill out questionnaires, and an unwillingness to undertake additional work as reasons for their disinterest. Moreover, many teachers lacked confidence in their language proficiency, which also deterred them from participating. Consequently, the limited number of participants undermined the generalizability of the study's findings.

5.7. Recommendations for Further Research

As this study was based on voluntary participation, the number of volunteers was limited. To improve the robustness and generalizability of the study's findings, it would be beneficial to replicate the study with a larger sample size. A larger sample would provide a more representative picture of the language teaching community and help ensure that the study's results are more reliable and accurate. Additionally, it may be helpful to incentivize participation by compensating participants or certifying their involvement with an institution. By doing so, more teachers may be willing to participate, leading to a more diverse and representative sample.

This would enhance the external validity of the study and improve its ability to provide insight into the professional development of language teachers.

In addition to the need for a larger sample size and incentivizing participation, the assessment literacy levels of the language teachers should be observed by evaluating their assessment practices both before and after the workshop. This would allow for a more comprehensive assessment of the impact of the course on the teachers' professional development. By examining the changes in assessment practices, it would be possible to determine if the workshop successfully improved the teachers' assessment literacy skills and if they were able to apply these skills to their classroom practices. This would provide valuable insights into the effectiveness of the workshop and could help guide the development of future professional development programs for language teachers.

Finally, to gain a more comprehensive understanding of the language assessment related needs of English language teachers across Turkey and other countries, this study should be replicated in different contexts. The professional development needs of language teachers may vary depending on the country, region, or even school setting. Therefore, it is important to examine the effectiveness of workshop for professional development in various contexts. Replicating the study in different settings could help identify specific areas of improvement for language teachers in each context and help develop more tailored professional development programs that cater to their unique needs. Additionally, conducting the study in other countries would provide a more global perspective on language assessment literacy and the effectiveness of workshop professional development programs for language teachers.

REFERENCES:

- Afshar, H. S. & Ranjbar, N. (2021). EAP Teachers' Assessment Literacy: From Theory to Practice. *Studies in Educational Evaluation*, (70). 101042. Retrieved from 10.1016/j.stueduc.2021.101042.
- Alonzo, D., Labad, V., Bejano, J., & Guerra, F. (2021). The Policy-driven Dimensions of Teacher Beliefs about Assessment. *Australian Journal of Teacher Education*, 46(3). Retrieved from <http://dx.doi.org/10.14221/ajte.2021v46n3.3>
- Arias, C., & Maturana, L. (2005). Evaluación en lenguas extranjeras: discursos y prácticas [Assessment in language teaching: Discourse and practices]. *Íkala, Revista de Lenguaje y Cultura*, 10(1) 63-91. Retrieved from <https://doi.org/10.17533/udea.ikala.3051>
- Ary, D., Jacobs, L.C. & Razavieh, A. (1996). *Introduction to Research in Education*. (6th Edition). Fort Worth: Harcourt brace College Publishers.
- Bachman, L. F., & Palmer, S. A. (1996). *Language Testing in Practice: Designing and Developing Useful Language Tests*. Oxford: Oxford University Press.
- Bachman, L. F., & Palmer, A. (2010). *Language Assessment in Practice: Developing Language Assessments and Justifying their use in the real World useful Language Tests*. Oxford: Oxford University Press
- Berry, V., Sheehan, S., & Munro, S. (2019). What does Language Assessment Literacy mean to teachers. *ELT Journal Volume* 73(2). Retrieved from 113-123. doi:10.1093/elt/ccy055
- Black, P., & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education*, 5(1), 7-73. Retrieved from <https://doi.org/10.1080/0969595980050102>
- Bogdan, R. C., & Biklen, S. K. (1998). *Qualitative Research for Education: An Introduction to Theory and Methods*. Ally & Bacon: Needham Heights, MA.
- Brown, D. H., (2004). *Language Assessment: Principles and Classroom Practices*. New York: Pearson Education.
- Brown, J. D., (1996). *Testing in Language Programs*. New Jersey: Prentice Hall Regents.

- Büyükkarcı, K. (2014). Assessment Beliefs and Practices of Language Teachers in Primary Education. *International Journal of Instruction*, 7(10), 107-120.
- Çakan, M. (2004). Öğretmenlerin Ölçme-Değerlendirme Uygulamaları ve Yeterlik Düzeyleri: İlk ve ortaöğretim. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, (37), 99–114.
- Çelik, S., Çelik, H., & Coombe, C. (Eds.) (2021). *Language Assessment and Test Preparation in English as a Foreign Language (EFL) Education*. Ankara: Vizetek Yayıncılık.
- Chan, Y. C. (2008). Elementary school EFL teachers' beliefs and practices of multiple assessments. *Reflections on English Language Teaching*, 7(1), 37-62
- Cheng, L., Rogers, T., & Hu, H. (2004). ESL=EFL Instructors' Classroom Assessment Practices: Purposes, Methods, and Procedures. *Language Testing* 21(3). 360-389.
- Coombe, C., Vafadar, H. & Mohebbi, H. (2020). Language Assessment Literacy: what do we need to learn, unlearn, and relearn? *Language Testing in Asia*, 10(3). 1-16.
- Creswell, J. W., (2014). *Research Design: Qualitative, Quantitative and Mixed Methods Approaches (4th Edition)*. SAGE Publications.
- Cui, Y., Yunlong, L., Gao, Y. (2022). Developing English teachers' Language Assessment Literacy in an EAP Reform Context through Test Design: A case study. *System* 109(1). 1-12.
- Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing*, 25(3), 327–347. <https://doi.org/10.1177/0265532208090156>
- DeLuca, C., & Volante, L. (2016). Assessment for learning in teacher education programs: Navigating the juxtaposition of theory and praxis. *Journal of the International Society for Teacher Education*, 20, 19–31.
- Deulen, A.A. (2013). Social Constructivism and Online Learning Environments: Toward a Theological Model For Christian Educators. *Christian Education Journal: Research on Educational Ministry*, (10)1, 90-98.
- Díaz, C., Alarcón, P., & Ortiz, M. (2012). El profesor de inglés: sus creencias sobre la evaluación de la lengua inglesa en los niveles primario, secundario y terciario [The English teacher: His beliefs about English language assessment at primary, secondary

- and tertiary levels]. *Íkala, Revista de Lenguaje y Cultura*, 17(1), 15-26. Retrieved from <https://doi.org/10.17533/udea.ikala.10039>
- Doğan, Ö. (2023). *Beyond the surface: An Investigation into the Language Assessment Literacy of EFL teachers working at K12 Schools*. Unpublished master's thesis, Yıldız Teknik University. İstanbul, Türkiye.
- Douglas, D. (2010). *Understanding Language Testing*. New York: Hodder Education.
- Esfandiari, R., & Nouri, R. (2016). A mixed methods, Cross-sectional Study of Assessment Literacy of Iranian University Instructors: *Implications for Teachers' Professional Development*, 19(2). 115-154.
- Farhady, H., & Tavassoli, K. (2018). Developing a Language Assessment Knowledge Test for EFL teachers: A data-driven approach. *Iranian Journal of Language Teaching Research*, 6(3), 79–94.
- Feretti, F., Santi, G. R. P., Del Zozzo, A., Garzetti, M., & Bolondi, G. (2021). Assessment practices and beliefs: Teachers' perspectives on assessment during long distance learning. *Education in Science*, 11(6). 264.
- Frodden, M., Restrepo, M., & Maturana, L. (2004). Analysis of assessment instruments used in foreign language teaching. *Íkala, Revista de Lenguaje y Cultura*, 9(1), 171- 201.
- Froozi, T., Razavipour, K. & Ahmadi, A. (2019). The language assessment literacy needs of Iranian EFL teachers with a focus on reformed assessment policies. *Language testing In Asia*, 9(2). 1-14. Retrieved from <https://doi.org/10.1186/s40468-019-0078-7>
- Fulcher, G. (2012). Assessment Literacy for the Language Classroom. *Language Assessment Quarterly*, 9, 113–132.
- Gelbal, S., & Kelecioğlu, H. (2007). Öğretmenlerin Ölçme ve Değerlendirme Yöntemleri Hakkındaki Yeterlikleri ve Karşılaştıkları Sorunlar. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 33, 135-145.
- Giraldo, F. (2018). Language Assessment Literacy: Implications for Language Teachers. *Profile Issues in Teachers' Professional Development*, 20(1), 179-195. Retrieved from <https://doi.org/10.15446/profile.v20n1.62089>
- Giraldo, F. (2021). A reflection on initiatives for teachers' professional development through language assessment literacy. *Profile Issues in Teachers' Professional Development*, 23(1), 197-213. Retrieved from <https://doi.org/10.15446/profile.v23n1.83094>

- Giraldo, F., & Murcia, D. (2018). Language Assessment Literacy for Pre-Service Teachers: Course Expectations from Different Stakeholders. *Gist Education and Learning Research Journal*, 16, 56-77.
- Gönen, K. & Akbarov, A. (2015). Instructors' Principles and Practices of Classroom Based Language Assessment in Higher Education in Turkey. *Journal of European Education*, 5(3), 28-38.
- Graham, P. (2005). Classroom-Based Assessment: Changing Knowledge and Practice through Preservice Teacher Education. *Teaching and Teacher Education*, 21(6), 607-621.
- Guerin, E. M. C. (2010). Initial findings from a pilot Italian study of foreign language teachers' stated language assessment knowledge-base and needs. Paper presented at the Lancaster University Postgraduate Conference in Linguistics & Language Teaching, Lancaster, the United Kingdom.
- Hakim, B. (2015). English Language Teachers' Ideology Of ELT Assessment Literacy. *International Journal of Education & Literacy Studies*, 3(4), 42-48.
- Han, T. & Kaya, H. İ. (2014). Turkish EFL Teachers' Assessment Preferences And Practices In The Context Of Constructivist Instruction. *Journal of Studies in Education*, 4(1), 77-93.
- Harding, L., & Kremmel, B. (2016). "Teacher assessment literacy and professional development," In D. Tsagari & J. Banerjee (Eds.), *Handbook of Second Language Assessment*. (Berlin, Boston, MA: De Gruyter), 413–428.
- Hasselgreen, A., Carlsen, C., & Helness, H. (2004). *European Survey of Language Testing and Assessment Needs*. Retrieved from www.ealta.eu.org/resources.htm
- Hatch, E., & Lazaraton, A. (1991). *The Research Manual: Design and Statistics for Applied Linguistics*. Boston: Heinle & Heinle Publisher.
- 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. (2015a). English Language Testing and Evaluation (ELTE) Training In Turkey: Expectations and Needs of Pre-Service English Language Teachers. *ELT Research Journal*, 4(2), 111-128.

- Hatipoğlu, Ç. (2015b). Diversity in language testing and assessment literacy of language teachers in Turkey. Paper presented at the 3 rd ULEAD Congress, International Congress on Applied Linguistics: Current Issues in Applied Linguistics, Çanakkale, Turkey.
- Hatipoğlu, Ç., & Erçetin, G. (2016). Türkiye’de Yabancı Dilde Ölçme ve Değerlendirme Eğitiminin Geçmişi ve Bugünü. In S. Akcan & Y. Bayyurt (Eds), 3. *Ulusal Yabancı Dil Eğitimi Kurultayı: Türkiye’deki Yabancı Dil Eğitimi Üzerine Görüş Ve Düşünceler* 23-24 Ekim 2014, konferanstan seçkiler, 72-89. İstanbul, Turkey: Boğaziçi University.
- Herrera, L., & Macías, D. (2015). A Call for Language Assessment Literacy in the Education and Development of Teachers of English as a Foreign Language. *Colombian Applied Linguistics Journal*, 17(2), 302-312. Retrieved from <https://doi.org/10.14483/udistrital.jour.calj.2015.2.a09>.
- Hidri, S. (2015). Conceptions of Assessment: Investigating What Assessment Means to Secondary and University Teachers. *Arab Journal of Applied Linguistics*, 1(1), 19-43.
- Hill, K. (2017). Understanding classroom-based assessment practices: a precondition for teacher assessment literacy. *Papers in Language Testing and Assessment*, 6(1). 1-17.
- Hughes, A., (2003). *Testing for Language Teachers*. Cambridge: Cambridge University Press.
- Inbar-Lourie, O. (2008). Constructing a Language Assessment Knowledge Base: A Focus on Language Assessment Courses. *Language Testing*, 25(3), 385 – 402.
- Inbar-Lourie, O. (2013a). Guest editorial to the special issue on language assessment literacy. *Language Testing*, 30(3), 301–307. <https://doi.org/10.1177/0265532213480126>
- Inbar-Lourie, O. (2013b, November). *Language assessment literacy: What are the ingredients?* [Paper presentation]. 4th cbla sig Symposium, University of Cyprus, Cyprus
- Ivankova, N. V., Creswell, J. W., & Stick, S. L. (2006). Using Mixed-Methods Sequential Explanatory Design: From Theory To Practice. *Field Methods*, 18(1), 3-20.
- Janatifar, M., & Marandi, S. (2018). Iranian EFL Teachers' Language Assessment Literacy (LAL) under an Assessing Lens. *Applied Research on English Language*.7(3), 361-382.

- Jannati, S. (2015). ELT Teachers' Language Assessment Literacy: Perceptions and Practices. *The International Journal of Research in Teacher Education*, 6(2), 26-37.
- Jeong, H. (2013). Defining Assessment Literacy: Is It Different for Language Testers and Non-Language Testers? *Language Testing*, 30(3), 345-362. Retrieved from <https://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. Retrieved from <https://doi.org/10.1177/0265532209351431>.
- Johnson, M. W. (2016). The Personal Learning Environment and The Institution of Education: Reflections on Technological Personalisation In Itec Schools. *International Journal of Smart Technology and Learning* 1(1), 93.
- Johnson, D. W., & Johnson, R. T. (1999) Making Cooperative Learning Work. *Theory Into Practice*, 38(2), 67-73.
- Jones, I. S., & Blankenship, D. (2017). Learning Style Preferences and The Online Classroom. *Research in Higher Education Journal*, 33(8).
- Karaman, P. & Şahin, Ç. (2014). Öğretmen Adaylarının Ölçme Değerlendirme Okuryazarlıklarının Belirlenmesi. *Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi (KEFAD)*, 15(2), 175-189.
- Karin V., Dina, T. & Georgios Spanoudis (2020) What Do Teachers Think They Want? A Comparative Study of In-Service Language Teachers' Beliefs on LAL Training Needs, *Language Assessment Quarterly*, 17(4), 386-409. Retrieved from DOI: 10.1080/15434303.2020.1781128
- Kavaklı, N., & 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), 1261-1270.
- Kırkgöz, Y. & Ağçam, R. (2012). Investigating the Written Assessment Practices of Turkish Teachers of English at Primary Education. *The Journal of Language and Linguistic Studies*, 8(2), 119-136.
- Koh, K. H. (2011). Improving Teachers' Assessment Literacy through Professional Development. *Teaching Education*, 22(3), 255-276.

- Koh, K., Burke, L. E. C., Luke A., Gong, W. & Tan, C. (2017). Developing the assessment literacy of teachers in Chinese language classrooms: A focus on assessment task design. *Language Teaching Research*, 22(3), 1-25.
- Köksal, D. (2004). Assessing teachers' testing skills in ELT and enhancing their professional development through distance learning on the net. *Turkish Online Journal of Distance Education-TOJDE*, 5(1).
Retrieved from <https://dergipark.org.tr/tr/pub/tojde/issue/16931/176755>
- Kremmel, B. & Harding, J. (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.
Retrieved from DOI: 10.1080/15434303.2019.1674855ö
- Lam, R. (2015). Language Assessment Training in Hong Kong: Implications for Language Assessment Literacy. *Language Testing*, 32(2), 169-197. Retrieved from doi:10.1177/026553221455432
- Lan, C. & Fan, S. (2019). Developing Classroom-Based Language Assessment Literacy For In-Service EFL Teachers: The Gaps. *Studies in Educational Evaluation*, (61), 112-122
- Latif, M. W., & Wasim, A. (2022). Teacher Beliefs, Personal Theories and Conceptions of Assessment Literacy-A Tertiary EFL Perspective. *Language Testing in Asia*, 12(11), 1-22. Retrieved from <https://doi.org/10.1186/s40468-022-00158-5>
- Leap, K., Channy, M., & Chan, N. K. (2015). Cambodian ELT University Practitioners' Use of Standardized Tests For Practice and Assessment. *Language Education in Asia*, 6(1), 4-16.
- Levi, T., & Inbar-Lourie, O. (2020). Assessment Literacy or Language Assessment Literacy: Learning from the Teachers. *Language Assessment Quarterly*, 17(2), 168–182.
Retrieved from <https://doi.org/10.1080/15434303.2019.1692347>.
- Lian, L. H., Yew, W. T., & Meng, C. C. (2014). Enhancing Malaysian Teachers' Assessment Literacy. *International Education Studies*, 7(10), 74–81. Retrieved from <https://doi.org/10.5539/ies.v7n10p74>.
- López, A., & Bernal, R. (2009). Language Testing in Colombia: A Call for more Teacher Education and Teacher Training in Language Assessment. *Profile: Issues in Teachers' Professional Development*, 11(2), 55-70.

- Mahapatra, S. K. (2016). Tracking the Development of Teachers' Language Assessment Literacy: A Case Study. In G. Pickering & P. Gunashekar (Eds.), *Ensuring quality in English Language Teacher Education*, 106-114. New Delhi, India: British Council.
- Malone, M. E. (2013). The Essentials of Assessment Literacy: Contrasts Between Testers and Users. *Language Testing*, 30(3), 329–344.
- Mede, E., & Atay, D. (2017). English Language Teachers' Assessment Literacy: The Turkish Context. *Ankara Üniversitesi TÖMER Dil Dergisi*, 168(1), 43-6.
- Mertler, C. A., & Campbell, C. (2005). Measuring teachers' knowledge & application of classroom assessment concepts: Development of the Assessment Literacy Inventory. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Quebec, Canada.
- Ministry of National Education (MoNE). (2017). *Öğretmenler için Genel Yeterlikler*. Retrieved from https://oygm.meb.gov.tr/meb_iys_dosyalar/2018_06/29111119_TeachersGeneralCompetencies.pdf
- Ministry of National Education. (2017). *General Competencies for Teaching Profession*.
- Montee, M., Bach, A., Donovan, A., & Thompson, L. (2013, Spring). LCTL teachers' assessment knowledge and practices: An exploratory study. *Journal of the National Council of Less Commonly Taught Languages*, 14, 1-32.
- Muhammad, F. H. N. & Bardakçı, M. (2019). Iraqi EFL Teachers' Assessment Literacy: Perceptions and Practices. *Arab World English Journal*, 10 (2) 420-430. Retrieved from <https://dx.doi.org/10.24093/awej/vol10no2.33>
- Muñoz, A., Palacio, M., & Escobar, L. (2012). Teachers' Beliefs About Assessment in an EFL Context in Colombia. *Profile: Issues in Teachers' Professional Development*, 14(1), 143-158.
- Ölmezer-Öztürk, E., & Aydın, B. (2019). Investigating Language Assessment Knowledge of EFL Teachers. *Hacettepe University Journal of Education*, 34(3), 602-620. Retrieved from doi:10.16986/HUJE.2018043465
- Öz, H. (2014). Turkish Teachers' Practices for Assessment for Learning in the English as a Foreign Language Classroom. *Journal of Language Teaching and Research*, 5(4), 775-785.

- Öz, S. & Atay, D. (2017). Turkish EFL instructors' in-class language assessment literacy: Perceptions and practices. *ELT Research Journal*, 6(1), 25-44.
- Palincsar, A. S. (1998). Social constructivist perspectives on teaching and learning. *Annual Review of Psychology*, 49, 345–375. Retrieved from <https://doi.org/10.1146/annurev.psych.49.1.345>
- Pastore, S., & Andrade, H. L. (2019). Teacher Assessment Literacy: A three dimensional model. *Teaching and Teacher Education*, 84, 128-138.
- Perry, M. (2013). *Teacher and Principal Assessment Literacy*. Doctoral Dissertation, University of Montana. Retrieved from <https://scholarworks.umt.edu/cgi/viewcontent.cgi?article=2410&context=etd>
- Pill, J., & Harding, L. (2013). Defining the language assessment literacy gap: Evidence from a parliamentary inquiry. *Language Testing*, 30(3), 381 – 402
- Plake, B. S., Impara, J. C., & Fager, J. J. (1993). Assessment Competencies of Teachers: A National Survey. *Educational Measurement: Issues and Practice*, 12(4), 10-12. Retrieved from <http://dx.doi.org/10.1111/j.1745-3992.1993.tb00548>
- Popham, W. J. (2009). Assessment Literacy For Teachers: Faddish Or Fundamental? *Theory into Practice*, 48(1), 4-11. Retrieved from <https://doi.org/10.1080/00405840802577536>
- Popham, W. J. (2011). Assessment literacy overlooked: A teacher educator's confession. *The Teacher Educator*, 46(4), 265–273. Retrieved from <https://doi.org/10.3390/educsci11060264>.
- Riestenberg, K., Silvio, F. D., Donovan, A., & Malone, M. E. (2010). Development of a computer-based workshop to foster language assessment literacy. *Journal of the National Council of Less Commonly Taught Languages*, 9, 21-42.
- Roberts, J. (1998). *Language Teacher Education*. London: Arnold, A member of Hodder Headline Group.
- Rogers, W. T., Cheng, L., & Hu, H. (2007). ESL/EFL Instructors' Beliefs About Assessment And Evaluation. *Canadian and International Education/Education canadienne at internationale*, 36(1), 39-61.
- Şahin, S. (2019). *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.
- Sariyildiz, G. (2018). *A study into language assessment literacy of pre-service English as a foreign language teachers in Turkish context* Unpublished master's thesis, Hacettepe University.
- Sariyildiz-Canlı, G. (2023). *Language Assessment Literacy Level and Needs of In-Service EFL Teachers in Turkey*. Unpublished doctoral dissertation, Hacettepe University.
- Sellan, R. (2017). Developing Language Assessment Literacy in Singapore: How Teachers Broaden English Language Learning by Expanding Assessment Constructs. *Papers in Language Testing and Assessment*, 6(1), 64-87.
- Semiz, Ö. & Odabaş, K. (2016). Turkish EFL teachers' familiarity of and perceived needs for language testing and assessment literacy. Proceeding of the 3rd international linguistics and language studies conference, İstanbul, Turkey. Retrieved from https://docs.wixstatic.com/ugd/bac820_2c683311d65a478f9b10ee6dfbd1020d.p d
- Shohamy, E., Inbar-Lourie, O., & Poehner, M. (2008). *Investigating assessment perceptions and practices in the advanced foreign language classroom* (Report No. 1108). University Park, PA: Center for Advanced Language Proficiency Education and Research.
- Sultana, N. (2019). Language Assessment Literacy: An Uncharted Area for the English Language Teachers in Bangladesh. *Language Testing in Asia*, 9(1), 1-14. Retrieved from <https://doi.org/10.1186/s/40468-019-0077-8>
- Sun, H. & Zhang, J. (2022) Assessment Literacy of college EFL teachers in China: Status quo and mediating factors. *Studies in Educational Evaluation*, 74.
- Tajeddin, Z., Khatib, M., & Mahdavi, M. (2022). Critical language assessment litreacy of EFL teachers: Scale Construction and Validation. *Language Testing*, 39(4), 649-678.
- Talib, R., Kamsah, M. Z., Ghafar, M. N. A., Zakaria, M. A. Z. M., & Naim, H. A. (2013). Tassess: Assessment literacy for Malaysian teachers. Paper presented at the International Conference on Assessment for Higher Education Across Domains and Skills, Kuala Lumpur.
- Taylor, L. (2009). Developing Assessment Literacy. *Annual Review of Applied Linguistics*, 29, 21-36.

- Taylor, L. (2013). Communicating the theory, practice and principles of language testing to test stakeholders: Some reflections. *Language Testing*, 30(3), 403-412.
- Vogt, K., & Tsagari, D. (2014). Assessment Literacy of Foreign Language Teachers: Findings of a European Study. *Language Assessment Quarterly*, 11(4), 374-402.
Retrieved from: DOI: 10.1080/15434303.2014.960046
- Vogt, K., & Tsagari, D. (2017). *Enhancing the Language Assessment Literacy of teachers – A study of training needs across Europe*. Presented at the 14th Annual Conference of EALTA, Center international d'études pedagogiques, Serves, France. June 1-3, 2017.
- Vogt, K., Tsagari, D., Csépes, I., Green, A., & Sifakis, N. (2020). Linking learners' perspectives on language assessment practices to teachers' assessment literacy enhancement (TALE): Insights from four European countries. *Language Assessment Quarterly*, 17(4), 410–433. Retrieved from <https://doi.org/10.1080/15434303.2020.1776714>.
- Volante, L., & Fazio, X. (2007). Exploring teacher candidates' assessment literacy: Implications for teacher education reform and professional development. *Canadian Journal of Education*, 30, 749–770.
- Xu, Y., & Brown, G. T. L. (2017). University English Teacher Assessment Literacy: A survey-test report from China. *Papers in Language Testing and Assessment* 6(1), 133-158.
- Xu, Y., & Liu, Y. (2009). Teacher Assessment Knowledge and Practice: A Narrative Inquiry of a Chinese College EFL Teacher's Experience. *Tesol Quarterly*, 43(3).
- Yan, X. & Fan, J. (2020). Am I qualified to be a language tester? Understanding the development of language assessment literacy across three stakeholder groups? *Language Testing*, 38(1), 1-28.
- Yan, X. & Zhang, C. & 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).
- Yapıcı, M. ve C. Demirdelen (2007). İlköğretim 4. sınıf programına ilişkin öğretmen görüşleri. *İlköğretim Online*, 6(2), 204-212.
- Yastıbaş, A. E. (2018). *Language Assessment Literacy Of Turkish EFL Instructors: A Multiple-Case Study*. Unpublished Doctoral Dissertation, Atatürk University.

Yetkin, C. (2015). *An Investigation on ELT Teacher Candidates' Assessment Literacy*

Unpublished master's thesis, Çağ University.

Yüksek Öğretim Kurumu. (2007). *Öğretmen Yetiştirme ve Eğitim Fakülteleri (1982-2007)*.

Retrieved from <https://www.yok.gov.tr/Documents/Yayinlar/Yayinlarimiz/ogretmen-yetistirme-ve-egitim-fakulteleri.pdf>



APPENDIX A: INFORMED CONSENT FROM

APPENDIX B: ETHIC COMMITTEE REPORT



APPENDIX C: PROFESSIONAL KNOWLEDGE OF ASSESSMENT QUESTIONNAIRE

Dear Respondent,

The purpose of this questionnaire is to collect information on the English language teachers' professional knowledge of language assessment. The information you provide may lead to major changes in teacher education programs. Therefore, your careful answers will help researchers to have valid data. The information will be kept strictly confidential and will be used just for research purposes. We will also send a summary of our findings to participants upon request.

Name (optional): _____

Email address (optional): _____

Phone number (optional): _____

Q1. What is your gender? Female Male

Q2. What is your age range?

- Under 20
- 21 - 25
- 26 - 30
- 31 - 35
- 36 - 40
- 41 - 45
- 46 - 50
- 51 - 55
- 56 - 60
- Above 60

Q3. How many years of teaching experience (public schools, private institutes, tutoring) do you have?

Q4. What is your educational degree?

- High School Graduate
- Associate degree
- BA degree
- MA degree
- Doctorate
- Other

Q5. Where is your hometown?

Assessing Writing														
Assessing Speaking														
Formative Assessment														
Summative Assessment														
Diagnostic Assessment														
Peer and Self-Assessment														
Performance Assessment														
Portfolio Assessment														
Giving Feedback to Students														
Developing Fair and Objective Scoring Rubric														
Item Analysis (IF, ID, CD)														
Reliability of The Test														
Validity of The Test														
Interpreting Test Scores														
Knowledge of Statistics														
Ethical Considerations in Testing														

Q13. Do you have any other comments that will help us to understand your language assessment-related needs in an online language testing course?

APPENDIX D: PROFESSIONAL KNOWLEDGE OF ASSESSMENT (PKA) TEST

Part A.

For questions 1-5, match the situations with the appropriate test types listed A-F.

Mark the correct letter (A-F) on your answer sheet.

There is one extra option.

Situations

1. The English department of a university administers an English test at the end of the BA program to check the overall English knowledge of its students.
2. The parents want to make a wise decision on what their child should study in the future.
3. A language institute is planning to hire a number of teachers. Only teachers who obtain the predetermined criterion score on an English test will be hired.
4. A language institute administers an English test to all incoming students to make sure they are directed into appropriate classes.
5. A language teacher uses a test at the end of the course to check how much students have learned from the materials taught during the course.

Test Types

- A. Achievement
- B. Aptitude
- C. Diagnostic
- D. Placement
- E. Proficiency
- F. Selection

Part B.

For questions 6-10, put the stages of language test design in order.

Mark the correct letter (B-F) on your answer sheet.

The first stage (A) is done for you. You do not need to use option A again.

0	A	Stages of Language Test Design
6	___	A. Deciding on the function and form of the test
7	___	B. Validating the test to determine test characteristics
8	___	C. Planning for the content of the test based on table of specifications
9	___	D. Asking a colleague to give comments on the test
10	___	E. Writing different types of items in accordance with the specified content
		F. Pretesting the items with a group similar to the target group to determine item characteristics

Part C.

For questions 11-15, read the scenario. Choose which statistical technique, listed A-F, is used by the teacher in doing each task.

Mark the correct letter (A-F) on your answer sheet.

There is one extra option.

Statistical Techniques

- | |
|-------------------------------|
| A. Choice Distribution |
| B. KR-21 Reliability |
| C. Item Discrimination |
| D. Item Facility |
| E. Pearson Correlation |
| F. Point Biserial Correlation |

A language teacher developed a final exam for her class. In doing so, she followed the steps of test design carefully. In the pretesting stage, she administered the test to a group similar to her students' level. After collecting the data, she did the following tasks to explore the characteristics of individual items and the whole test.

Tasks

11. She divided the number of correct responses to each item by total number of responses to that item.
12. She divided the students into two groups of high and low ability to see how well each item differentiated between them.
13. She found the percentage of students selecting each of the alternatives in multiple-choice items.
14. She checked the consistency of total scores by putting the number of items, and mean and variance of the test in a formula.
15. She found the degree of correspondence between the scores from her test and a criterion test, which were administered together.

Part D.

For questions 16-23, choose the correct letter (A, B, or C) to complete the sentences.

Mark the correct letter (A, B, or C) on your answer sheet.

16. A language teacher developed a final exam for his class. He tried to reduce the amount of potential error in his measurement. This is an issue related to test _____.
A. practicality
B. reliability
C. validity
17. A teacher developed a new test. Later on, he was asked by his supervisor, "Does this test measure the ability we want to investigate?" Such a question is asked when one is concerned with _____ validity.
A. concurrent
B. consequential
C. construct
18. A variable is defined as a characteristic that changes from person to person, object to object, place to place, or time to time. The process of quantifying a variable through using an instrument is called _____.
A. evaluation
B. measurement
C. testing
19. A group of testers was asked to develop a series of achievement tests to measure well-defined and specific objectives of a program, and to assess the amount of each student's learning. Such tests are considered _____.
A. criterion-referenced
B. norm-referenced
C. performance-based
20. Standardized tests are developed by a group of testing experts. In these tests, _____.
A. test takers are differentiated by their ranks
B. items have predetermined responses
C. certain objectives are followed

21. A language teacher piloted his newly developed test in a class. The measurement scales needed to code the participants' gender and test scores are _____ respectively.
- A. nominal and ordinal
 - B. nominal and interval
 - C. ordinal and interval
22. The tests mostly associated with the integrative approach in language testing are _____ tests.
- A. cloze and dictation
 - B. translation and essay-type
 - C. true-false and multiple-choice
23. A would-be language tester learned in a training session that there are different types of reliability, including _____.
- A. content
 - B. predictive
 - C. rater

Part E.

A language teacher is using different assessment activities in his class to evaluate his students' progress during the course better.

For questions 24-28, match the classroom assessment activities he uses with their corresponding assessment techniques, listed A-F.

Mark the correct letter (A-F) on your answer sheet.

There is one extra option.

Assessment Techniques

- A. Conferences
- B. Journals
- C. Observations
- D. Self-Assessment
- E. Peer-Assessment
- F. Portfolios

Classroom Assessment Activities

- 24. He frequently uses one-to-one interactions with individual students to give feedback to them, to provide them with positive washback, and to check their progress.
- 25. He keeps a record of his students' works including the different compositions they have written during the course, their projects, and results of their tests. This way, he will have multiple sources of information to evaluate his students.
- 26. He pays attention to his students' behavior and performance in the class through every question and response, and keeps a record of these for later scoring.
- 27. He often asks his students to work together, to give comments to each other, and to evaluate each other's work impartially.
- 28. He asks his students to write about their thoughts, feelings, reactions, ideas, or progress towards course goals regularly. Through these, he can check their improvement throughout the course.

Part F.

A language teacher is planning to test different language components/skills in her class. To do so, there are certain issues she should pay attention to.

For questions 29-35, choose the correct letter (A, B, or C) to complete the sentences.

Mark the correct letter (A, B, or C) on your answer sheet.

29. In a diagnostic test of grammar, _____.
A. each item should test one grammatical point
B. the choices should be of different sizes
C. the stem should represent a simple sentence
30. To test pronunciation at elementary levels, it is better to use _____.
A. retelling
B. intonation recognition
C. sound discrimination
31. One of the most effective ways to determine what words to include in a general vocabulary test is to use the _____ of words.
A. frequency
B. stability
C. scope
32. To test listening, one can use an activity in which noise is added to the background. This activity is called _____.
A. dictation
B. distorted message
C. listening cloze
33. To select a passage for a test of reading, one should ensure that the passage is _____, and at the appropriate difficulty level for the test takers.
A. suitable, seen
B. reliable, suitable
C. unseen, authentic

34. Depending on the purpose and evaluation criteria of interviews, the teacher assigns a rating such as excellent/ good/ fair/ poor. This marking system is called ____.
- A. holistic
 - B. discrete
 - C. analytic
35. To avoid subjectivity in scoring the compositions, the teacher decides to score them in two different occasions. In this case, she is mostly concerned about ____ reliability.
- A. inter-rater
 - B. intra-rater
 - C. test-retest

APPENDIX E: PKAQ ITEM ALIGNMENT TABLE

PKAQ Item 10-11-12	Related Factor	CFA for Item 10	CFA for Item 11	CFA for Item 12
Item 1 (Procedures in Language Test Design)	2 (Basics of test Development)	2	2	2
Item 2 (Deciding on What To test)	2 (Basics of test Development)	2	2	2
Item 3 (Writing Test Specifications)	2 (Basics of test Development)	2	2	2
Item 4 (Selecting Items for the Test)	2 (Basics of test Development)	2	2	2
Item 5 (Preparing Items)	2 (Basics of test Development)	2	2	2
Item 6 (Test Administration)	2 (Basics of test Development)	4	2	2
Item 7 (Assessing Grammar)	3 (Assessing Language Skills)	4	1	3
Item 8 (Assessing Pronunciation)	3 (Assessing Language Skills)	3	1	1
Item 9 (Assessing Vocabulary)	3 (Assessing Language Skills)	3	1	3
Item 10 (Assessing reading Comprehension)	3 (Assessing Language Skills)	3	1	1
Item 11 (Assessing Listening)	3 (Assessing Language Skills)	3	1	1
Item 12 (Assessing Writing)	3 (Assessing Language Skills)	3	1	1
Item 13 (Assessing Speaking)	3 (Assessing Language Skills)	3	1	1
Item 14 (Formative Assessment)	4 (Functions of Assessment)	2	1	2
Item 15 (Summative Assessment)	4 (Functions of Assessment)	3	4	3
Item 16 (Diagnostic Assessment)	4 (Functions of Assessment)	2	4	3
Item 17 (Peer and Self Assessment)	4 (Functions of Assessment)	2	1	1
Item 18 (Performance Assessment)	4 (Functions of Assessment)	2	1	1
Item 19 (Portfolio Assessment)	4 (Functions of Assessment)	2	1	1

Item 20 (Giving Feedback to Students)	2 (Basics of test Development)	1	1	1
Item 21 (Developing Fair and Objective Scoring Rubric)	2 (Basics of test Development)	1	3	3
Item 22 (Item Analysis) IF, ID,CD	1(Psychometric Features of Test)	1	3	4
Item 23 (Realibility of the Test)	1(Psychometric Features of Test)	1	3	4
Item 24 (Validity of the Test)	1 (Psychometric Features of Test)	1	3	4
Item 25 (Interpreting Test Scores)	1 (Psychometric Features of Test)	1	1	3
Item 26 (Kowledge of Statistics)	1 (Psychometric Features of Test)	1	3	4
Item 27 (Ethical Considerations in Testing)	2 (Basics of test Development)	1	3	3

APPENDIX F: THEORETICAL FRAMEWORK OF LAL COMPONENTS

Basics of Testing	✓ Procedures In Language Test Design ✓ Deciding on What to Test ✓ Writing Test Specifications/Blueprints ✓ Selecting Items for the Test ✓ Preparing Items ✓ Test Administration ✓ Giving Feedback to Students ✓ Developing Fair and Objective Scoring Rubric ✓ Ethical Considerations in Testing
Functions of Assessment	✓ Formative Assessment ✓ Summative Assessment ✓ Diagnostic Assessment ✓ Peer and Self-Assessment ✓ Performance Assessment ✓ Portfolio Assessment
Psychometric Features of A Test	✓ Item Analysis (IF, ID, CD) ✓ Reliability of The Test ✓ Validity of The Test ✓ Interpreting Test Scores ✓ Knowledge of Statistics
Assessing Language Skills	✓ Assessing Grammar ✓ Assessing Pronunciation ✓ Assessing Vocabulary ✓ Assessing Reading Comprehension ✓ Assessing Listening ✓ Assessing Writing ✓ Assessing Speaking

APPENDIX G: SUGGESTIONS

Dear Participants,

Thank you so much for being volunteer in this study. Below, you may find three questions to evaluate and express your suggestions about the sessions you have taken for three weeks.

Write down three things that you love most about the course:

-
-
-

Write down three things that you did not like about the course:

-
-
-

Write down one different thing that you have encountered througout the course. Explain whether you liked it or not.

.....

Anything that you want to suggest:

.....

APPENDIX H: PROMPTS FOR REFLECTIVE NOTES

Basics of Test Development

- Now... You are ready to design your own test. Why don't you give it a try?
- As a language teacher, you are supposed to prepare a language test that aims to measure the language ability of learners on making predictions.
- What are the steps you would follow? Explain Briefly.
- What is the test task you would use for the given situation?

Functions of Assessment

- Consider the type of decisions you are supposed to make during your teaching?
- What type of assessment do you use?
- What are some points you have discovered regarding assessment types at the end of this video?

Assessing Reading

- Consider the types of Reading Assessment Tasks you are using. What are they?
- Are there any types of assessment you want to share with me? If yes, please mention about it briefly.

Assessing Writing

- I would like you to write a short paragraph regarding your experience in teaching and assessing writing in your classes.
- Do you mind emphasizing more about the scoring methods you use?
- What are new points you loved most about assessing writing at the end of this course?

Assessing Listening

- Describe the challenges that you face while conducting listening tasks in your classes?

- Despite the challenges that you face while assessing listening, what are some assessment tasks that you use?
- Choose an appropriate type of assessment task that you would like to experience in your class while assessing listening. Briefly explain the procedure that you would follow.

Assessing Speaking

- Please reflect on your experiences regarding Assessment of Speaking. What are some tasks that you also prefer using in your classes?
- What are some challenges that you face?
- What are some points that you have just decided to consider during the assessment of speaking in your teaching?

Revision

Please reflect on the videos you have watched so far.

APPENDIX I: WORKSHOP SYLLABUS

Date	Lessons	Topics	References
May 11-13	Basics of Test Development	DESIGN Target Language Use Sample Task Analysis OPERATIONALIZATION Test Task Specification Writing Instructions Blueprint Scoring Method ADMINISTRATION Try-Out Operational use Procedures Ethical Considerations	Bachman, L. F., & Palmer, S. A., (1996) Language testing in Practice: Designing and Developing Useful Language Tests. Oxford University Press.
May 16-20	Functions of Assessment	Informal vs Formal Assessment Assessment Types and Thinking Skills Summative vs Formative Assessment NRT vs CRT	Bachman, L. F., & Palmer, A., (2010) Language Assessment in Practice: Developing Language Assessments and Justifying their use in the real World useful

Characteristics of NRT and CRT Language Tests. Oxford
University Press

Alternative Language Assessment

Peer and Self Assessment

Portfolio Assessment
(Collection, Reflection and
Selection)

Dynamic Assessment

Task based Assessment

Performance Assessment

Projects

Language Test Types

Achievement Tests

Proficiency Tests

Diagnostic Tests

Aptitude Tests

Brown, D. H., (2004)
Language Assessment:
Principles and Classroom
Practices. Longman

May 23-27

Psychometric
Features of a
Test

PART 1

Principles of testing , Evaluation and Assessment in Foreign Language Teaching

Authenticity

Reliability

- Student-related
Reliability

- Rater-Related Reliability
- Test- related Reliability
- Administration related reliability
- More about Reliability

Brown, J. D., (1996)
Testing in Language
Programs: Prentice Hall
Regents

Validity

- Content Validity
- Criterion Validity
- Construct Validity
- Consequential Validity
- Face Validity

Practicality

Douglas, D., (2010)
Understanding Language
Testing. Hodder Education

Backwash

Security

Transparency

PART 2

Item Analysis

A. NRT

- IF (Item Facility)
- ID (Item Discrimination)

B. CRT

Item Content Analysis

Item Quality Analysis

DI

B-Index

Hatch , E., & Lazaraton,
A., (1991) The Research
Manual: Design and
Statistics for Applied
Linguistics. Heinle &
Heinle Publisher

Scales of Measurement

- Nominal
- Ordinal
- Continuous

Displaying Data

Describing Test Results

- ✓ Central Tendency

Mean, Mode, Median, Midpoint

- ✓ Dispersion

Range, standard deviation, variance

Hughes, A., (2003) Testing for Language Teachers. Cambridge University Press.

Normal Distribution and Skewedness

May 30-June 10 Assessing Language Skills

Assessing Speaking

- ✓ Imitative Assessment Tasks
- ✓ Intensive Assessment Tasks
- ✓ Responsive Assessment Tasks
- ✓ Extensive Assessment Tasks
- ✓ Interactive Assessment Tasks

Assessing Writing

- ✓ Designing Tasks for Testing Writing
- ✓ Sample Classroom Tasks
- ✓ Scoring and Rater Training

Assessing Reading

- ✓ MultipleChoice
- ✓ Cloze test
- ✓ ShortAnswer
- ✓ Summary
- ✓ Information Transfer Task
- ✓ Integrated Skills
- ✓ Scoring Reading AssessmentTasks
- ✓ Text Difficulty and Readability

Coombe, C., Çelik, S& Çelik, H. (Ed.). (2021). Language Assessmentand Test Preparation in English as Foreign Language (EFL) Education. Vizetek Yayıncılık

Assessing Listening

- ✓ Intensive Assessment Tasks
- ✓ Selective Assessment Tasks
- ✓ Responsive Assessment Tasks
- ✓ Extensive Assessment Tasks

June 13-17

Revision

Which Item to Choose?**IF vs ID****DE: Distractor Efficiency****Displaying Data (Mean,
Mode,Median, Range,
Midpoint)****Descriptive Statistics****Reliability****Central Tendency vs Dispersion****Global/Holistic Scales****Analytic Scales****Criterion Referenced Scales**

APPENDIX J: PKAQ CFAS WITH 4 FACTOR SOLUTION

(Item 10)Rotated Component Matrix^a

	Component			
	1	2	3	4
1.Procedures in Language Test Design		,687		
2. Deciding on What to Test		,710		
3. Writing Test Specifications/Blueprint		,673		
4.Selecting Items for the Test		,713		
5. Preparing Items		,692		,389
6. Test Administration				,626
7. Assessing Grammar				,800
8. Assessing Pronunciation			,498	,470
9. Assessing Vocabulary		,366	,466	
10. Assessing Reading Comprehension	,489		,579	
11. Assessing Listening	,371		,564	
12. Assessing Writing	,376		,699	
13. Assessing Speaking			,870	
14.Formativ Assessment		,513		
15.Summative Assessment		,465	,566	
16. Diagnostic Assessment		,634		
17. Peer and Self Assessment		,562	,362	
18. Performance Assessment		,646	,430	
19. Portfolio Assessment	,444	,473	,360	
20. Giving Feedback to Students	,618	,511		
21. Developing Fair and Objective Scoring Rubric	,562			
22. Item Analysis (IF,ID,CD)	,786			,365

23. Reliability of the Test	,888			
24. Validity of the Test	,887			
25. Interpreting Test Scores	,612		,417	
26. Knowledge of Statistics	,554	,367		
27. Ethical Considerations in Testing	,631		,443	



(Item 11) Rotated Component Matrix^a

	Component			
	1	2	3	4
1.Procedures in Language Test Design		,667		
2. Deciding on What to Test		,749		
3. Writing Test Specifications/Blueprint		,841		
4.Selecting Items for the Test		,825		
5. Preparing Items	,475	,623		
6. Test Administration	,498	,615		
7. Assessing Grammar	,702			
8. Assessing Pronunciation		,644		
9. Assessing Vocabulary	,732			
10. Assessing Reading Comprehension	,801			
11. Assessing Listening	,677			
12. Assessing Writing	,747			
13. Assessing Speaking	,691			
14.Formativ Assessment	,501			,457
15.Summative Assessment				,864
16. Diagnostic Assessment				,804
17. Peer and Self Assessment	,577	,377		
18. Performance Assessment	,646	,422		
19. Portfolio Assessment	,486	,409		,405
20. Giving Feedback to Students	,558			,362
21. Developing Fair and Objective Scoring Rubric			,513	
22. Item Analysis (IF,ID,CD)			,829	
23. Reliability of the Test			,939	
24. Validity of the Test			,920	
25. Interpreting Test Scores	,501		,482	
26. Knowledge of Statistics			,594	
27. Ethical Considerations in Testing	,368	,357	,408	

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 6 iterations.

(Item 12)Rotated Component Matrix^a

	Component			
	1	2	3	4
1.Procedures in Language Test Design		,693		
2. Deciding on What to Test		,737	,359	
3. Writing Test Specifications/Blueprint		,772		
4.Selecting Items for the Test		,816		
5. Preparing Items		,772		
6. Test Administration	,409	,570		
7. Assessing Grammar	,385		,507	
8. Assessing Pronunciation	,694			,385
9. Assessing Vocabulary	,434		,567	
10. Assessing Reading Comprehension	,707		,380	
11. Assessing Listening	,811			
12. Assessing Writing	,690		,409	
13. Assessing Speaking	,801			
14.Formativ Assessment	,360	,456		
15.Summative Assessment			,728	
16. Diagnostic Assessment			,704	
17. Peer and Self Assessment	,633		,387	
18. Performance Assessment	,608	,486		
19. Portfolio Assessment	,602			
20. Giving Feedback to Students	,407	,367		
21. Developing Fair and Objective Scoring Rubric			,612	
22. Item Analysis (IF,ID,CD)				,765
23. Reliability of the Test				,900
24. Validity of the Test				,902

25. Interpreting Test Scores			,706	
26. Knowledge of Statistics	,454		,362	,570
27. Ethical Considerations in Testing		,386	,547	

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 8 iterations.

