

INVESTIGATING THE ASSESSMENT APPROACHES OF EFL TEACHERS IN
A PRIVATE SCHOOL IN ANKARA: IMPACT OF TEACHING EXPERIENCE

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF SOCIAL SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

GÖZDENUR ÇETİN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF ARTS
IN
THE DEPARTMENT OF ENGLISH LANGUAGE TEACHING

SEPTEMBER 2025

Approval of the thesis:

**INVESTIGATING THE ASSESSMENT APPROACHES OF EFL TEACHERS
IN A PRIVATE SCHOOL IN ANKARA: IMPACT OF TEACHING
EXPERIENCE**

submitted by **GÖZDENUR ÇETİN** in partial fulfillment of the requirements for the degree of **Master of Arts in English Language Teaching, the Graduate School of Social Sciences of Middle East Technical University** by,

Prof. Dr. Sadettin KIRAZCI
Dean
Graduate School of Social Sciences

Prof. Dr. Nurten BİRLİK
Head of Department
Department of Foreign Language Education

Assist. Prof. Dr. Müge GÜNDÜZ
Supervisor
Department of Foreign Language Education

Examining Committee Members:

Prof. Dr. Cemal ÇAKIR (Head of the examining committee)
Gazi University
Department of English Language Teaching

Assist. Prof. Dr. Müge GÜNDÜZ (Supervisor)
Middle East Technical University
Department of Foreign Language Education

Prof. Dr. Çiler HATİPOĞLU
Middle East Technical University
Department of Foreign Language Education

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: Gözdenur ÇETİN

Signature:

ABSTRACT

INVESTIGATING THE ASSESSMENT APPROACHES OF EFL TEACHERS IN A PRIVATE SCHOOL IN ANKARA: IMPACT OF TEACHING EXPERIENCE

ÇETİN, Gözdenur

M.A., The Department of English Language Teaching

Supervisor: Assist. Prof. Dr. Müge GÜNDÜZ

September 2025, 141 pages

The main purposes of the current study are to investigate the assessment approaches of in-service EFL teachers working at a private school in Ankara, to examine how teaching experience affects their assessment orientations, and to explore how these teachers perceive classroom assessment and reflect on their practices. The study employs a case study design with an embedded mixed methods approach. 18 teachers took part in the study, and the data were collected through an adapted version of Approaches to Classroom Assessment Inventory (ACAI) and semi-structured interviews. In-service teachers' assessment approaches were identified based on the 4 key domains of the ACAI framework (DeLuca et. al, 2016), namely assessment purpose, assessment process, assessment fairness and assessment theory. Quantitative data were analyzed through descriptive analysis, and qualitative data were analyzed through the inductive analysis approach. Quantitative findings revealed that the assessment for learning approach (AfL) in the assessment purpose theme, the design approach in the assessment process theme and the consistent theory approach in the assessment theory theme received the highest level of support from the participants. Support for the assessment fairness approaches was relatively low across all

scenarios. In addition, moderately experienced teachers showed higher support for most assessment themes compared to their more experienced colleagues. Qualitative data revealed that the participants viewed assessment as a tool to keep a record of student progress, guide instructional activities and support learning. Considering these findings, the study hopes to draw attention to aspects which may inform professional development activities and programs and institutional policies.

Keywords: In-service EFL teachers, classroom assessment, assessment approaches, teaching experience

ÖZ

ANKARA'DAKİ BİR ÖZEL OKULDA GÖREV YAPAN İNGİLİZCE ÖĞRETMENLERİNİN ÖLÇME-DEĞERLENDİRME YAKLAŞIMLARININ İNCELENMESİ: ÖĞRETMENLİK DENEYİMİNİN ETKİSİ

ÇETİN, Gözdenur

Yüksek Lisans, İngiliz Dili Öğretimi Bölümü

Tez Yöneticisi: Dr. Öğr. Üyesi Müge GÜNDÜZ

Eylül 2025, 141 sayfa

Bu çalışmanın temel amaçları Ankara'daki bir özel okulda görev yapan İngilizce öğretmenlerinin ölçme değerlendirme yaklaşımlarını incelemek, öğretmenlik deneyiminin bu yaklaşımlar üzerindeki etkisini araştırmak ve öğretmenlerin sınıf içi ölçme değerlendirmeyi nasıl algıladıklarını ve bunu pratiğe nasıl yansıttıklarını araştırmaktır. Bu çalışma gömülü karma yöntem yaklaşımıyla yapılandırılmış bir durum çalışmasıdır. Çalışmaya 18 öğretmen katılmıştır ve veriler Sınıf İçi Ölçme ve Değerlendirme Yaklaşımları Envanteri'nin (ACAI) uyarlanmış bir versiyonu ve yarı yapılandırılmış görüşmeler aracılığıyla toplanmıştır. Öğretmenlerin ölçme değerlendirme yaklaşımları ACAI çerçevesinde belirtilen 4 ana tema olan ölçme ve değerlendirmenin amacı, süreci, adaleti ve kuramı temaları üzerinden belirlenmiştir. Nicel veriler betimsel analiz yoluyla nitel veriler ise tümevarımsal analiz yaklaşımıyla analiz edilmiştir. Nicel veriler, değerlendirmenin amacı temasında öğrenme için değerlendirme yaklaşımının, süreç temasında tasarıma dayalı yaklaşımın ve kuram temasında tutarlı kuram yaklaşımının katılımcılar tarafından daha fazla desteklendiğini ortaya koymuştur. Ölçme değerlendirmede adalet

temasında ise tüm yaklaşımlar bütün senaryolarda az düzeyde destek almıştır. Ayrıca, orta düzeyde deneyime sahip öğretmenler tüm ölçme değerlendirme temalarına daha fazla deneyime sahip olan öğretmenlere nazaran daha fazla destek göstermiştir. Bu bulgular doğrultusunda, çalışma mesleki gelişim etkinlikleri ve programları ile kurumsal politikalara ışık tutabilecek yönlerde dikkat çekmeyi amaçlamaktadır.

Anahtar kelimeler: İngilizce öğretmenleri, sınıf içi ölçme ve değerlendirme, ölçme ve değerlendirme yaklaşımları, öğretmenlik deneyimi

To my beloved family

ACKNOWLEDGEMENTS

First of all, I would like to express my deepest gratitude to my thesis advisor, Asist. Prof. Dr. Müge GÜNDÜZ for her invaluable guidance, support and constructive feedback throughout the research. Her encouragement, kindness and patience helped me keep going forward in this process.

I am also sincerely grateful to the members of the examining committee, Prof. Dr. Çiler HATİPOĞLU and Prof. Dr. Cemal ÇAKIR, for their valuable time, insightful comments and inspiring feedback.

I would like to express my heartfelt gratitude to my family, who have always been my greatest supporters throughout my life. I am deeply indebted to my mother Endam ÇETİN, whose love, guidance, and encouragement lifted me up whenever I doubted myself. I am equally thankful to my father Nazım Şeref ÇETİN and my brother Mert ÇETİN, whose patience, understanding, and constant support have been a source of strength for me throughout this journey.

I am especially grateful for my dearest friend Emel Nur ER, who has stood by me at every stage of my life, through joy and tough times. I am also thankful to my dear friends Büşra UYSALBAŞ DAVARCIOĞLU, Özlem Duru SAĞLAM and Esmâ TOKATLI, whose support and faith in me have been invaluable throughout this process.

Finally, I would also like to express my sincere thanks to my colleagues who kindly took the time from their busy schedules to participate in my research. This research would not have been possible without their willingness and contribution.

TABLE OF CONTENTS

PLAGIARISM	iii
ABSTRACT	iv
ÖZ.....	vi
DEDICATION	viii
ACKNOWLEDGEMENTS	ix
TABLE OF CONTENTS	x
LIST OF TABLES	xiii
LIST OF ABBREVIATIONS	xiv
CHAPTERS	
1. INTRODUCTION.....	1
1.1. Background to the study	1
1.2. Statement of the problem.....	3
1.3. Purpose of the study and research questions	4
1.4. Significance of the study	5
1.5 Definition of important terms	6
2. LITERATURE REVIEW.....	8
2.1. Assessment	8
2.2. Classroom assessment	11
2.3. Assessment Literacy (AL).....	13
2.4. Language Assessment Literacy (LAL).....	16
2.5. Tools and Models Used in TAL and LAL Research	20
2.6. Studies Conducted on TAL and LAL.....	23
2.6.1. International Studies Conducted on TAL and LAL.....	23
2.6.2. Studies Conducted on TAL and LAL in Türkiye	32
3. METHODOLOGY	41
3.1. Research design	41
3.2. Research setting	42

3.3. Participants	45
3.4. Data collection.....	47
3.4.1. Quantitative data collection instrument and data collection process	48
3.4.2. Qualitative data collection instrument and data collection procedure	52
3.5. Data analysis.....	53
3.6 Trustworthiness	54
3.7. The researcher’s position.....	55
4. RESULTS	57
4.1. In-service EFL Teachers’ Approaches to Assessment	57
4.2. Teaching Experience and Support for Assessment Approaches	70
4.3. In-service EFL Teachers’ Perceptions and Experiences Regarding Classroom Assessment	73
4.3.1 Perceptions and Experiences Regarding Assessment Purpose and Use	74
4.3.2. Perceptions and Experiences Regarding Assessment Practices	78
4.3.3. Perceptions and Experiences Regarding Assessment Fairness.....	87
4.3.4. Perceptions and Experiences Regarding Assessment Challenges	90
4.3.5. Perceptions and Experiences Regarding Professional Development	93
5. DISCUSSION	96
5.1. Discussion	96
6. CONCLUSION	105
6.1 Conclusion.....	105
6.2. Implications of the study	106
6.3. Limitations of the study and recommendations for further research.....	108
REFERENCES.....	109
APPENDICES	
A. APPROVAL OF THE METU HUMAN SUBJECTS ETHICS COMMITTEE	120
B. ADAPTED VERSION OF THE ACAI	121
C. INTERVIEW QUESTIONS IN ENGLISH	126
D. INTERVIEW QUESTIONS IN TURKISH	127
E. INFORMED CONSENT FORM.....	128
F. THE CODEBOOK	130

G. TURKISH SUMMARY	131
H. THESIS PERMISSION FORM / TEZ İZİN FORMU.....	141

LIST OF TABLES

Table 3.1	Demographic profile of the participants	46
Table 3.2	Assessment themes and approaches in ACAI	48
Table 4.1	Descriptive statistics for Scenario 1	58
Table 4.2	Descriptive statistics for Scenario 2	60
Table 4.3	Descriptive statistics for Scenario 3	62
Table 4.4	Descriptive statistics for Scenario 4	65
Table 4.5	Descriptive statistics across 4 scenarios for all participants	68
Table 4.6	Participants' assessment approaches	69
Table 4.7	Comparison of MET and HET	71
Table 4.8	Codes derived from in-service EFL teachers interview responses	74

LIST OF ABBREVIATIONS

EFL	English as a Foreign Language
AL	Assessment Literacy
LAL	Language Assessment Literacy
TAL	Teacher Assessment Literacy
ACAI	Approaches to Classroom Assessment Inventory
AoL	Assessment of Learning
AfL	Assessment for Learning
AaL	Assessment as Learning
MET	Moderately Experienced Teachers
HET	Highly Experienced Teachers
TALQ	Teacher Assessment Literacy Questionnaire
API	Assessment Practices Inventory
CALI	Classroom Assessment Literacy Inventory
LAKS	Language Assessment Knowledge Scale
IEP	Individualized Education Plan
CEFR	Common European Framework of Reference
ENLTA	European Network of Language Testing and Assessment
PD	Professional Development

CHAPTER 1

INTRODUCTION

1.1. Background to the study

Assessment has always been a central part of classroom practice. Teachers use assessment for many different purposes, such as monitoring students' progress, guiding instruction and providing feedback to students and other stakeholders (Phye, 1996; Popham, 2013). Teachers need to have a deeper understanding of assessment because effective assessment enables them to support learning outcomes and make meaningful instructional decisions. Therefore, teacher assessment literacy has attracted increasing attention both in research and practice over the last twenty years (DeLuca et al., 2016a).

In the late 19th century, educational measurement and psychometrics led to the formation of modern assessment approaches and standardized testing (Stiggins, 1991). Summative assessment started to dominate teaching practices by serving various functions such as dividing students into proficiency groups, measuring achievement and ensuring validity and reliability in assessing students. However, by the mid-20th century, limitations of summative assessment were recognized by researchers and educators. Formative assessment practices were introduced by scholars such as Scriven and Bloom, and they highlighted that formative assessment could guide and inform instructional practices (Black & Wiliam, 1998). Then, Black and Wiliam (1998) emphasized the importance of formative assessment in enhancing student learning. Following all these developments in the field, many alternative assessment methods such as peer and self-assessment, projects and portfolios were introduced to have a more comprehensive understanding of student learning. Formative and alternative assessments have gained attention from both

educators and researchers ever since, and it has been thoroughly studied in the literature. However, despite the attempts in research to emphasize the importance of assessment, it is still not used effectively to improve learning in primary and middle school classrooms (Tierney, 2006). The emphasis on summative assessment is still strong, and everyday assessment practices stay largely the same (Tierney, 2006). This resistance to change is also evident in the Turkish context. Even though there is awareness of classroom assessment to some extent, there is a conflict between classroom practices and assessment literacy (Öz & Atay, 2017). This mismatch between assessment literacy and classroom practices can be due to many reasons. Yet, a part of it is due to the challenges that educators face. Tierney (2006) claims that shifting the assessment focus from traditional summative assessments to alternative and formative assessments is not an easy task to accomplish because educational changes are likely to be slow and gradual. Also, it requires preparation and readiness on the teachers' part. It was reported that the majority of the teachers found the assessment training provided at their institutions inadequate and did not feel ready to use alternative assessments to improve student learning (Tierney, 2006). In a similar regard, DeLuca et al., (2019) stated that many teachers, especially in the early years of their careers, felt unprepared to use assessment to support learning. Furthermore, the new variety of assessment tools and methods made it even more difficult for teachers to understand, implement and interpret these practices while they try to report evidence of student learning to all stakeholders (DeLuca et al., 2019). Pastore (2023) agrees with this view by stating that teachers still have difficulties with assessment practices especially when they try to integrate new tools or theories into their classrooms. Therefore, it can be said that teachers experience a lack of confidence regarding their assessment literacy levels and assessment practices. This lack of confidence is also rooted in the history of the field, as it did not get enough attention at the beginning in the field of teacher education (Pastore, 2023). In addition to a lack of confidence in assessment practices, one main challenge that teachers face is a lack of professional development (PD) activities and training in pre-service education. Vogt et al. (2008) stated that most of the teachers did not receive assessment training focusing on different types of assessment serving various purposes. The situation is not any different in the Turkish context. According to Ölmezer-Öztürk (2021), graduates of the teaching departments get only one course

on assessment during their pre-service education. Also, in Türkiye, graduates of literature and culture departments can become English teachers, and unfortunately, they do not take any courses related to assessment and evaluation (Büyükkarcı, 2016). Despite all challenges, assessment plays a key role in English language teaching as it supports learning, guides instructional practices and provides valuable feedback to all stakeholders. Therefore, it continues to attract attention in the literature, teacher education and professional development programs.

In addition to assessment literacy, teachers' approaches to classroom assessment, their assessment orientations and their perceptions regarding the role of assessment should also be taken into consideration. Thus, in this study, the four-dimensional ACAI framework developed by DeLuca et al. (2016) was used to better understand the assessment practices and approaches of in-service EFL teachers in a private middle school context. This framework views assessment literacy as a dynamic and context-bound system and organizes it around four key domains, namely assessment purpose, assessment process, assessment fairness and assessment theory. The assessment purpose domain focuses on choosing goal-oriented assessment practices. The assessment process domain examines how teachers design, implement and interpret assessments. The assessment fairness domain examines how teachers ensure fairness in their assessment practices. Lastly, the assessment theory domain refers to the conceptual and pedagogical factors that shape a teacher's understanding of assessment. With the help of these domains, the framework aims to provide a more comprehensive way of examining teachers' assessment approaches.

1.2. Statement of the problem

Even though AL has been extensively examined in the literature, research that focuses on teachers' assessment approaches, which refers to teachers' assessment orientations regarding the purpose, process, fairness and theory of assessment, remains limited especially in private school contexts in Türkiye. However, teachers' assessment approaches are a critical area of study because it demonstrates how teachers position assessment in their daily teaching activities, and it reflects how assessment is implemented in different contexts.

In the Turkish EFL context, most of the studies focus on the inadequacies of teachers' assessment education and professional development (Öz & Atay, 2017; Büyükkarcı, 2016). Yet, how in-service English teachers approach classroom assessment gains relatively lower attention. Also, the application of ACAI in Turkish EFL context is still scarce in Türkiye despite its ability to provide a comprehensive framework for investigating assessment orientations. Moreover, the effect of teaching experience on assessment literacy has been investigated in the Turkish context, but the effect of teaching experience on assessment approaches needs further attention.

In sum, the problem that is addressed in this study is the lack of research on in-service EFL teachers' classroom assessment approaches in Türkiye in private school contexts through the ACAI framework.

1.3. Purpose of the study and research questions

This research study aims to investigate the assessment approaches of in-service EFL teachers working at a private school in Ankara, to examine how teaching experience affects their assessment orientations and to explore how teachers perceive and experience classroom assessment.

The following three research questions were formed in accordance with the purposes stated above:

1. To what extent do in-service EFL teachers support a specific classroom assessment approach in relation to the four ACAI framework themes of assessment (assessment purpose, assessment process, assessment fairness and assessment theory)?
2. How does teaching experience shape in-service EFL teachers' support for a specific classroom assessment approach in relation to the four ACAI framework themes of assessment (assessment purpose, assessment process, assessment fairness and assessment theory)?
3. How do in-service EFL teachers perceive classroom assessment and reflect on their practices regarding it?

1.4. Significance of the study

Assessment is one of the most important elements of teaching and learning, and teachers spend a lot of time on tasks and responsibilities related to assessment (Stiggins, 1999). Teachers' assessment practices help improve instruction and learning (Coombs et al., 2018). Even though teachers recognize the importance of assessment in teaching and learning, they often feel inadequate in their language assessment literacy (LAL) and struggle to improve it (Berry et al., 2019; Ergül & Çetin, 2021). Many teachers do not feel adequate when it comes to evaluating their students' performance (Mertler, 2003). Teachers are held accountable for their students' learning, but they do not feel prepared to fulfill this responsibility (Mertler, 2003). This problem gets worse because teacher education programs do not provide the teachers with the necessary knowledge and skills to meet the demands that come with the developments in the field and various assessment related reforms. According to Hatipoğlu (2015), there is only one course in language teacher education programs dedicated to evaluation and testing in language teaching. Therefore, it can be concluded that teachers are not being prepared to deal with the challenges of assessment because they are not provided with an opportunity to learn how to deal with them (DeLuca & Klinger, 2010). For this reason, it is essential to increase awareness of classroom assessment and provide teachers with practical, classroom-friendly assessment approaches that they can implement effectively (Willis et al., 2013; Coombs et al., 2018).

This study aims to help address this issue by examining how in-service English teachers with different levels of experience and assessment orientations perceive and experience classroom assessment. Thus, it can be beneficial for various stakeholders. English teachers may benefit from the study by being encouraged to reflect more on their assessment practices. It may help school administrators and professional development program developers in designing more targeted and effective in-service teacher training opportunities. Also, teacher educators can benefit from the study because it can help them identify the gaps in teacher education programs in terms of assessment practices and assessment approaches and encourage them to include more practical and teacher and classroom friendly practices in pre-service teacher

education courses on assessment. Furthermore, the teachers' assessment orientations in the current research context can provide valuable insights into how assessment is practiced in private school contexts. Similar to the current research context, the majority of the private schools in Türkiye have top-down summative assessment policies (Yıldırım-Seheryeli & Gelbal, 2020). While the teachers' autonomy over summative assessment is limited, parent expectations and high socio-economic backgrounds of the students continue to put pressure on teachers to continue to implement contemporary assessment practices. Therefore, by investigating how teachers practice assessment in such a context, the study aims to demonstrate that assessment literacy is not only a professional competency, but it is also a concept that is greatly influenced by institutional demands, constraints and complex social dynamics. Moreover, to the best of the researcher's knowledge, there is limited research focusing on the assessment approaches of in-service EFL teachers specifically in private middle school contexts in Türkiye through the ACAI framework. This study aims to address this gap by offering an in-depth analysis of how institutional and personal factors shape classroom assessment and whether teaching experience plays a role in shaping teachers' assessment approaches. By examining EFL teachers' assessment approaches through the ACAI framework, this study hopes to contribute to the literature on classroom assessment approaches in private educational settings.

1.5 Definition of important terms

Assessment: A process through which teachers assess students' performance by gathering, analyzing and interpreting information related to a specific objective in the curriculum (Gnaicha, 2016).

Classroom assessment: A combination of formative and summative assessment techniques that are used to evaluate student performance, inform teaching activities and improve student learning (Shepard, 2019).

Assessment literacy: A concept that not only includes competencies such as designing, implementing, interpreting and communicating assessment practices but

also represents a developmental process that is shaped by cultural and educational context, desire and opportunity to learn, ethics and systemic expectations (Willis et al., 2013; DeLuca et al., 2016)

Language assessment literacy: “Language assessment literacy refers to the knowledge, skills and principles that stakeholders involved in assessment activities are required to master in order to perform assessment tasks” (Inbar-Lourie, 2017, p. 257).

Teacher assessment literacy: Skills and knowledge that teachers develop and improve through everyday practice to measure, evaluate and assess student achievements and support student learning (Xu & Brown, 2016).

Assessment approach: A teacher’s theoretical understanding and practical application of assessment within a specific educational setting (Coombs et al., 2018).

Assessment for learning (AfL): Helping students progress towards learning goals and providing evidence of learning to adapt teaching activities (Coombs et al., 2018).

Assessment of learning (AoL): Recording student learning and grading student performance in relation to students’ achievement of learning goals (Coombs et al., 2018).

Assessment as learning (AaL): An approach through which students actively use assessment to monitor, reflect on, and direct their own learning (Coombs et al., 2018).

CHAPTER 2

LITERATURE REVIEW

Chapter 2 presents the definition of assessment and classroom assessment, description of assessment types, the evolution and development of language assessment literacy (LAL), tools, frameworks and models used in LAL and TAL research, and studies conducted internationally and in Türkiye to examine teachers' assessment literacy.

2.1. Assessment

Assessment plays a key role in teaching and learning, and it has been defined in various ways in the literature. Brookhart (1999) defined assessment as collecting and analyzing information about student accomplishments. Black and Wiliam (2018) defined assessment as a process that allows educators to draw conclusions about student learning. Data regarding learning are generated through assessment practices and teachers use these data as evidence of learning. Therefore, it can be said that “Educational assessment is at heart an exercise in evidentiary reasoning” (Mislevy & Riconscente, 2005, p. iv). This view focuses on the conclusions drawn from the assessment rather than the assessment itself. Other definitions emphasize that assessment is more than keeping scores as it allows for an interpretation of student performance. Brown (2004) defined assessment as understanding and explaining the performance of students with the help of a variety of methods or practices. Expanding on this view, Gnaicha (2016) defined assessment as a process through which teachers assess students' performance by gathering, analyzing and interpreting information related to a specific objective in the curriculum. He also added that assessment required providing feedback both to the teachers and students based on the information gathered from the assessment, which highlighted the role of

assessment as a bridge between evaluation, teaching and learning practices. Therefore, assessment is not limited to theory, but it is also embedded in daily teaching and learning activities within the classroom.

In addition to the definitions, researchers emphasized the importance of assessment in supporting teaching and learning. Tosuncuoğlu (2018) stated that assessment played a key role in learning and teaching as teachers could evaluate their students' performance, identify their needs and adapt instruction accordingly based on the results of appropriate assessment. Black and Wiliam (2018) focused on another reason for the importance of assessment by highlighting the complex relationship between teaching and learning. They stated that assessment was crucial because what students learned and what they were taught did not always align with each other. Therefore, educators needed to develop strategies or methods that would allow them to elicit evidence of learning. Only then would they be able to adjust instruction and respond to students' needs to enhance their learning (Black & Wiliam, 2018).

Educators use multiple methods and techniques to identify the learning gaps. Summative and formative assessment are distinguished by researchers as two interconnected functions that enable educators to have a more comprehensive understanding of student learning. Formative assessment refers to the continuous evaluation process through which educators collect evidence of learning throughout the teaching process to adapt instruction in ways that would meet students' needs and keep track of their progress (Gnaicha, 2016). On the other hand, summative assessment refers to the evaluation of student performance at the end of an instructional unit (Black & Wiliam, 1998). This distinction is also evident in the widely recognized framework of assessment of learning (AoL), assessment for learning (AfL) and assessment as learning (AaL). AoL aligns closely with summative assessment as it refers to recording student learning and grading performance in relation to students' achievement of learning goals (Coombs et al., 2018). On the other hand, AfL aligns closely with formative assessment as it refers to helping students progress towards learning goals and providing evidence of learning to adapt teaching activities (Coombs et al., 2018). AaL is different from AfL and AoL as it places the learners at the center of the learning process by encouraging self-

monitoring and effective use of feedback to enhance learning (Coombs et al., 2018). In addition to these traditional functions, a variety of alternative methods have been introduced by researchers to broaden the concept of assessment. Alternative methods “such as journals, logs, portfolios, self-assessment, and peer-assessment help reveal “what students can do with language” and “they require students to perform, create, and produce something” (Phongsirikul, 2018, p. 62). In this way, the learners take ownership of the learning process. Therefore, alternative methods suggest that in addition to assessment’s role in identifying learning gaps, it also functions as a motivational tool and contributes to the quality of learning and teaching (Phongsirikul, 2018).

Teachers can collect evidence of learning by incorporating a variety of assessment types such as formative, summative and alternative assessment. This variety of assessments enables the teachers to gain a deeper understanding of student learning and make meaningful instructional decisions. As a result, assessment contributes to professional development and instructional improvement. From this perspective, it can be concluded that assessment has various benefits for teachers such as evaluating students’ knowledge and capabilities, determining areas for improvement, providing motivation and receiving meaningful feedback on student learning (Tosuncuoğlu, 2018). Moreover, it is important to note that teachers’ continuous engagement in assessment contributes to professional confidence and increases collaboration among colleagues. In the study that they conducted with middle school teachers, Black and Wiliam (2018) reported that teachers felt more confident in implementing assessments and had more ownership of the procedures and tools when they were actively involved in designing them. The participants also indicated that they had to meet to design summative assessments and these meetings sparked meaningful discussions that were part of teachers’ learning processes.

Even though assessment may have a negative connotation for students, the significance is evident from students’ perspectives as well. Assessment also functions as a motivational tool and contributes to the quality of learning and teaching because the students might also want to keep track of their progress and evaluate their performance and achievements in relation to their classmates (Phongsirikul, 2018).

In conclusion, assessment is a fundamental component of learning and teaching. It refers to collecting data regarding students' performance and interpreting it to support instructional decisions. Formative assessment refers to continuous evaluation of students' performance to support their learning, and summative assessment refers to measuring achievement at the end of an instructional period. Also, there are alternative methods which help learners participate in the process of learning more effectively and actively by requiring them to produce or create something as evidence of their learning. Using different methods or tools together enables teachers to get a more complete picture of students' learning, adapt instruction, support professional development and increase students' engagement and enhance their self-monitoring skills.

2.2. Classroom assessment

As previously defined, assessment is a systematic process of collecting data about student learning to monitor students' progress, make instructional decisions and evaluate students' performance. Classroom assessment is an aspect of assessment, and its main purpose is to support and enhance student learning. Shepard (2019) defined classroom assessment as a combination of formative and summative assessment techniques that were used to evaluate student performance, inform teaching activities and improve student learning. She claimed that classroom assessment should be closely connected to instructional processes and activities and should be based on current research in motivation, feedback, self-monitoring, curriculum and teacher development because its main aim was to support learning rather than just to measure (Shepard, 2019).

In a similar way, Price et al. (2014) stated that classroom assessment was a crucial element of contemporary education. They pointed out that scholars and educators used to view assessment as standardized measures and tests, but recent research has recognized assessment as classroom-based measures of student achievement that are required for effective teaching and learning environments. Therefore, Black and Wiliam (1998) emphasized that classroom assessment should be a continuous part of daily teaching activities and routines so that it could fulfill its purpose. Expanding on

this view, DeLuca and Klinger (2010) highlighted that teachers were required to use a range of assessment tools to assess students' learning and progress to support instructional activities. They also underlined that assessment should promote self-monitoring and student autonomy in terms of learning. Similarly, Chappuis and Stiggins (2002) stressed that assessment should not only assign grades but also motivate and inspire students, and this was exactly what the AfL model aimed to do. Therefore, it can be said that classroom assessment practices closely align with the AfL model.

There are a variety of tools and methods that teachers can utilize in classroom assessment. However, Price et al. (2014) proposed six tools and broad categories that can be used to enhance learning. "The six assessment tools and strategies are: (1) rubrics, (2) performance-based assessments (PBAs), (3) portfolios, (4) student self-assessment, (5) peer-assessment, and (6) student response systems (SRS)." (Price et al., 2014, p. 6). Rubrics serve both as a tool to measure student understanding and as a way of checking whether students have understood the criteria or not. PBAs require students to internalize the knowledge and use it in situations which they are not familiar with. Self-assessments help students identify their strengths and weaknesses and encourage them to work further on these. Peer assessment positions students as evaluators. Finally, SRS allow teachers to collect evidence of learning through technological devices and systems. The benefits of these tools and strategies have been highlighted by many scholars in the field. DiCarlo and Cooper (2014) commented that teachers who were able to use classroom assessment techniques effectively helped their learners achieve their goals and improve their performance in a contemporary learning environment. Similarly, Black and Wiliam (1998) stated that AfL was one of the most effective approaches to enhance learning. Additionally, Chappuis and Stiggins (2002) stressed that students felt a sense of responsibility for their own learning when they were required to regularly evaluate and measure their performance.

Researchers recommend that further research be done to ensure the effectiveness of classroom assessment. DiCarlo and Cooper (2014) claimed that alternative online classroom assessments should be explored so that students could be provided with

instant feedback to get a deeper understanding of their mistakes before they left the classroom. Furthermore, Shepard (2019) drew attention to the fact that “students learn more from written comments alone than from comments plus grades” (Black et al. 2003; Butler, 1988 as cited in Shepard, 2019, p. 191). Therefore, educators and other stakeholders should expand their understanding of classroom assessment, and the focus of assessment should be on AfL practices (Shepard, 2019; Price et al., 2014).

2.3. Assessment Literacy (AL)

The National Council on Measurement in Education, the National Education Association and the American Federation of Teachers (1990) issued the *Standards for Teacher Competence in Educational Assessment of Students*. It was one of the earliest widely recognized attempts to define AL. According to the standards, “Assessment is defined as the process of obtaining information that is used to make educational decisions about students, to give feedback to the student about his or her progress, strengths, and weaknesses, to judge instructional effectiveness and curricular adequacy, and to inform policy.” (AFT et al., 1990, p. 2). There are 7 principles established by the standards that serve as a guide for teachers in designing, selecting and implementing assessments. The teachers are required to be competent in the following areas: 1) selecting suitable assessment methods for teaching decisions, 2) creating assessment methods that are suitable for teaching decisions, 3) applying, grading and analyzing the results of assessments, 4) making teaching decisions by interpreting assessment results, 5) creating valid grading procedures that are based on student assessments, 6) communicating assessment results with parents, learners and other educators, 7) identifying unethical assessment practices and their uses (AFT et al., 1990). After the release of the standards, earlier definitions of AL focused more on the technical dimensions of assessment. In other words, assessment literacy was initially viewed as a practical and professional skill that teachers should have, and it emphasized the importance of test construction and psychometric principles (Stiggins, 1991 as cited in DeLuca et al., in *Frontiers in Education*, p. 2) Paterno (2001) defines AL as “the possession of knowledge about the basic principles of sound assessment practice, including terminology, the development and

use of assessment methodologies and techniques, familiarity with standards of quality in assessment and familiarity with alternative to traditional measurements of learning” (Paterno, 2001 as cited in Mertler, 2003, p. 10). Stiggins (1991) defines AL by looking at what it is like “not to have it” (p. 535). He states that teachers who are not assessment literate cannot properly design effective assessments to get valuable assessment data or use the assessment data meaningfully. They may be easily confused by the technical aspects of assessment and accept the data without questioning it, so they do not possess the knowledge and the ability to critically interpret the results. On the contrary, teachers who are assessment literate can differentiate between good and poor assessments, and they are familiar with various forms of assessment. They question what the assessment shows the students about targeted learning goals and how it may affect students. They prefer assessments that clearly reflect the intended learning outcomes (Stiggins, 1991). Similarly, Popham (2009) states that assessment literacy (AL) refers to a teacher’s competency in understanding and using assessment to support learning. It also refers to the knowledge of diverse assessment tools and procedures. Assessment literate teachers can make important decisions, give effective feedback and adapt instruction based on assessment outcomes. Therefore, AL is considered to be one of the most important attributes of a competent teacher (Popham, 2009).

Webb (2002) defines AL from a broader perspective that includes the stakeholders and the effects of assessment on them. According to Webb (2002), AL is understanding how to evaluate students’ knowledge and capabilities, how to analyze the assessment outcomes and how to apply this knowledge to improve both student learning and program efficiency. He also extends this definition by saying that AL is much more than familiarity with assessment tools and analyzing results because assessment literate educators are also aware of “how assessments and the consequences of assessments create incentives and disincentives to students, teachers, principals, and staff to perform at their best” (Webb, 2002, p. 3). They know that assessment can be used to inform and educate important stakeholders such as teachers, parents, students and administrators (Webb, 2002).

Fulcher (2012) offers a more comprehensive definition by portraying AL as an understanding of principles and concepts that create the foundation for practice such as “ethics and codes of practice” and having information, capabilities and skills to create, implement or evaluate “large scale standardized or classroom-based tests” (p. 125). Also, he adds that AL includes the ability to evaluate the function and effects of assessment on the society, institutions and individuals and “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” (Fulcher, 2012, p. 125).

Willis et al. (2013) define AL as a dynamic, social and context bound practice. They state that AL is shaped by teachers working together with their colleagues and students to support intended learning goals through “classroom and cultural knowledge” (p. 252). This approach reconceptualizes assessment literacy as a collaborative and culturally shaped practice rather than an individualized effort. DeLuca et al. (2016) present AL not only as a critical professional competency that includes designing effective assessment, applying assessments and interpreting assessment outcomes to make meaningful decisions regarding instructional practices, but also as a source that would facilitate professional development by combining formative and summative assessment and connecting these with the previously established summative assessment standards. With this view, they move AL away from a list of technical abilities and conceptualize it as a multifaceted competence that is continuously developed through reflective practices and pedagogical awareness.

Xu and Brown (2016) describe AL as an essential component of professionalism. They compiled over 100 studies on teacher assessment literacy and created a framework called “Teacher Assessment Literacy in Practice” (TALIP) (Xu & Brown, 2016).

Their framework consists of 6 components:

1. Knowledge base: It includes information about assessment tools, designing assessment, interpreting results, feedback, knowledge of ethics and peer and self-assessment.

2. Teacher conceptions of assessment: It refers to teachers' beliefs, attitudes and values regarding assessment
3. Contextual effects: It refers to institutional, cultural and policy related effects
4. TALIP: It refers to the application of knowledge, beliefs and contextual factors into practice
5. Teacher learning: It refers to learning through reflective practices and PD
6. Assessor identity: It refers to a teacher's identity becoming a part of his/her professional self.

According to this framework, teachers show progress through a “continuum of mastery” (Xu & Brown, 2016, p. 159). They move from a basic knowledge of assessment principles and practices to establishing a connection between learning, teaching and assessment to design and use context-appropriate assessments and critically reflecting on them (Xu & Brown, 2016).

All of these definitions show that even though AL was defined based on its technical aspect at the beginning, the definitions evolved over time to include a much more comprehensive definition of AL that would capture it as a multidimensional concept including ethical, cultural and social aspects.

2.4. Language Assessment Literacy (LAL)

Language assessment literacy (LAL) has been recognized as an important aspect of effective language teaching all around the world. Therefore, there have been a lot of attempts to define it in a way that would be comprehensible for all stakeholders. Berry (2019) stated that in a broader sense, AL meant the knowledge that instructors should have about assessment. However, this term was later modified and adopted by experts in language education to reflect the complex nature of assessment in the field of language teaching. LAL is different from AL because it mainly focuses on assessment within language education. LAL focuses on aspects that are unique to language assessment because language has its own complexities and competencies (Weng & Shen, 2022).

Malone (2013) defines LAL as the ability of stakeholders, especially language teachers, to apply their understanding of assessment to classroom procedures in general and to language-related assessment concerns specifically. She also talks about important components of LAL that language teachers lack such as distinguishing between summative and formative assessment, using various assessment tools in language assessment and being aware of the different types of outcomes provided by them, giving effective feedback, promoting student autonomy, dealing with ethical issues such as the role of the assessor and the use of assessment outcomes, validity, reliability and fairness. Similarly, Weng and Shen (2022) emphasize that LAL is a crucial component of teachers' professional competence, which provides them with the means of designing, applying and analyzing assessment to enhance learning and make meaningful decisions regarding curriculum and lesson activities. Pill and Harding (2013) expand this view of LAL being an essential part of teachers' professional competence and state that they view LAL as "a repertoire of competences that enable an individual to understand, evaluate and, in some cases, create language tests and analyze test data" (p. 382).

Davies (2008) states that skills provide the training of essential and required methods, knowledge provides a pertinent context to language assessment, and principles offer insights into many aspects of assessment such as the correct use, fairness and effectiveness of assessment and issues related to ethics and professionalism. According to Davies (2008), a combination of skills and methodology is not sufficient; and a combination of skills and knowledge is not sufficient either without the addition of principles. Just like in learning, teaching also requires a balance of the practical and theoretical aspects because without one the other might be misunderstood or underestimated (Davies, 2008). This view has impacted the way many researchers conceptualize LAL today. In fact, it has been created by the historical development of the field as the shift in the last 40 years seems to be from a combination of skills and knowledge to a combination of skills, knowledge and principles (Davies, 2008).

Similarly, Inbar-Lourie (2008) also views LAL as a multifaceted concept combining convergent literacies, and she emphasizes that language-specific competencies

should be integrated into the LAL framework. She states that “The language assessment knowledge base in fact comprises layers of assessment literacy skills combined with language-specific competencies, forming a distinct entity that can be referred to as ‘language assessment literacy’ (henceforth, LAL).” (Inbar-Lourie, 2008, p.390). She later elaborated on this view by associating it with meaningful and critical engagement with assessment. She notes that being literate means having the capability to understand and question the purpose, the appropriacy and the conditions of an assessment (Inbar-Lourie, 2008, p. 389). In addition to seeing assessment as a multifaceted concept, Inbar-Lourie’s framework (2008) argues that language assessment should be seen as a “socially constructed activity” integrated into the “local context” where all stakeholders are acknowledged as significant partners in the assessment process (p. 386). This point of view defines assessment as a socially situated practice rather than an independent process. It encourages teachers to deal with issues related to decision-making processes by keeping the effects of the assessment on students in mind. While conducting assessments in this framework, it is crucial to think about planned and unplanned impacts assessment might have on particular groups and individuals and understand that assessment is “a social practice and a social product” (Filer, 2000, as cited in Inbar-Lourie, 2008, p. 387). This critical definition of assessment led to what Roever and McNamara (2006) called the “social turn” that happened in the field in the last two decades. The social turn suggests making fundamental changes related to conceptual development, offers critical views on the societal role of assessment and emphasizes the need to make assessment inclusive and fair and the accountability of the test makers (Lynch, 2001; Roever & McNamara, 2006).

Xu and Brown (2016) present an important reconceptualization of LAL, which is built on previous sociocultural perspectives. They define teacher assessment literacy (TAL) as something that teachers develop and improve through everyday practice rather than knowledge and a set of skills. They argue that LAL is shaped by a teacher’s beliefs, feelings, experiences and their educational context. Xu and Brown (2016) view teachers as “active agents who construct their understanding of assessment through participating in assessment practices within specific sociocultural contexts” (p. 150). In this framework, assessment is not only a technical process, but

it is also a sociocultural endeavor which requires constant critical thinking. They also note that assessment is not just “a matter of technical rationality”, but it also includes “moral and emotional dimensions” (Xu & Brown, 2016, p. 150). Xu and Brown’s (2016) understanding of LAL surpasses the earlier definitions of LAL to include what teachers do, know, believe, feel and adapt in relation to their educational contexts.

More recent research investigated how teachers’ philosophy and theoretical understanding affected their practice and conceptualization of assessment. This is represented in research as “assessment mindset”. DeLuca, Coombs and LaPointe-McEwan (2019) claimed that assessment mindset should be added to assessment literacy frameworks. The authors claimed that “assessment mindset is a component of assessment literacy and demarcates the specific linkage between teachers’ mindsets toward learning and their approach to assessment.” (DeLuca et al., 2019, p. 169). According to DeLuca et al. (2019), the mindset does not only inform a teacher's philosophy of learning, but it also influences the judgments teachers make about what effective assessment is, how fairness is defined, and whether teachers use assessment for learning or grading. Therefore, LAL includes more than a teacher’s skills and knowledge regarding assessment. It also includes teachers’ personal approaches and mindsets towards learning (DeLuca et al., 2019).

In conclusion, LAL is presented as a complex and multidimensional concept in the literature. It combines technical skills, theoretical knowledge, ethical principles, sociocultural awareness, and teachers’ personal beliefs and mindsets. Even though the definitions may vary, most of them agree that LAL focuses on teachers’ knowledge and capabilities and how these are shaped in different educational contexts. For the purposes of this study, LAL is defined as a concept that not only includes competencies such as designing, implementing, interpreting and communicating assessment practices but also as a developmental process that is shaped by cultural and educational context, desire and opportunity to learn, ethics and systemic expectations (Willis et al., 2013; DeLuca et al., 2016). The context-bound dimension of this definition is rooted in Willis, Adie and Klenowski’s (2013) sociocultural conceptualization of assessment. As explained before, they see AL as a

dynamic social practice through which teachers share their knowledge of classroom practices and culture with their students and colleagues to achieve targeted learning goals (Willis et al., 2013). The competencies-based part of the definition is rooted in the ACAI framework developed by DeLuca, LaPointe-McEwan and Luhanga in 2016. This framework was developed by analyzing “15 national and international assessment standards from six geographic regions (United States, Canada, United Kingdom, Europe, Australia, and New Zealand)” (DeLuca et al., 2016, p.250). The ACAI framework reflects the current changes in the field of assessment literacy such as integrating formative and summative purposes, focusing on diversity and fairness and the incorporation of assessment into sociocultural contexts. When Willis et al.’s (2013) sociocultural foundation of AL and ACAI’s competency framework are combined, they provide a strong base for the current study that enables an organized examination of how teachers’ perceptions, approaches, and context shape their assessment orientations.

2.5. Tools and Models Used in TAL and LAL Research

Many tools and models have been designed to examine teachers’ assessment literacy, but one of the earliest versions was “Teachers’ Assessment Literacy Questionnaire” (TALQ) (Plake et al., 1993). This tool was prepared to respond to teachers’ increasing concern regarding their readiness in this area (Plake et al., 1993). Research studies showed that even though teachers spent the majority of their time dealing with assessment related issues and responsibilities, they were inadequately prepared when it came to classroom assessment requirements, and this feeling was generally associated with “insufficient training in teacher education programs”. (Stiggins, 1991 & Plake et al., 1993). In order to deal with this issue, a tool was developed based on the Standards for Teacher Competence in Educational Assessment of Students (AFT et al., 1990) “to measure teachers’ competency levels in addressed standards” (Plake et al, 1993, p. 10). The tool was made up of two parts. Teachers’ knowledge of competency areas was measured with the help of 35 items in the first part of the questionnaire. In part two, teachers were asked questions about their educational background and their views on assessment. It was first administered to a group of in-service teachers and administrators in 1991. The results

revealed that the teachers were competent in designing, administering and interpreting the results of an assessment. However, they got a low score in communicating the assessment results. TALQ drew a lot of attention in the field since it was one of the earliest empirical tools to properly measure teachers' assessment literacy. Also, it was officially validated and grounded in nationally approved standards (Plake et al., 1993).

Another tool that was developed to be used in LAL research was the "Assessment Practices Inventory (API)", which was created by Zhang and Burry-Stock (1997). Unlike TALQ, API was not designed to measure the competency level of teachers in assessment areas, but it aimed to investigate the impact of assessment education and career stages on teachers' self-reported assessment competency (Zhang & Burry-Stock, 1997). There were 67 items, and the teachers were asked to evaluate their competencies from "Not at all skilled" to "Highly skilled" on a 5-point scale (p. 3). The instrument covered a wide range of assessment areas such as test design, interpretation of results, ethics and communication. The API was first administered to 311 teachers, and the teachers with assessment training reported higher levels of skill across items. API was crucial and highly valued in the literature because it paved the way for perception-based assessment literacy tools which could guide the design of teacher training and professional development programs (Zhang & Burry-Stock, 1997).

"Classroom Assessment Literacy Inventory (CALI)" was also developed to investigate teachers' understanding of assessment (Mertler, 2002). It aimed to measure knowledge-based competencies regarding assessment rather than self-perceptions. It was quite similar to TALQ. CALI was also aligned with the "Standards for Teacher Competence in the Educational Assessment of Students". It consisted of 35 multiple-choice and 7 background questions. The answers given to questions were scored and calculated, and this total score determined how knowledgeable teachers were about assessment related topics (Mertler, 2002). The first study with CALI was conducted with 67 preservice and 197 in-service teachers. The results revealed that in-service teachers scored higher than preservice teachers

(Mertler, 2002). It was later on used in many other studies because it provided a measure of factual and applied knowledge about teachers' assessment literacy.

After many studies were conducted using the tools explained above, a gap in LAL research was determined. According to DeLuca et al. (2016), technical knowledge has been the focus of traditional approaches to assessment literacy. However, recent research suggests that teachers' assessment literacy needs to take into account their methods, beliefs, and approaches within their context because AL cannot be fully understood without also examining teachers' methods, orientations, and particularly their assessment approaches within their contexts (DeLuca et al., 2016). Therefore, DeLuca, LaPointe-McEwan, and Luhanga developed the "Approaches to Classroom Assessment Inventory (ACAI)" in 2015 to focus more on how teachers understand and apply classroom assessment. The instrument measures four key dimensions of assessment approaches, which are assessment purpose, assessment process, fairness, and measurement theory (DeLuca et al., 2016, p. 249). It consists of three parts. In the first part, the participants are asked to answer some background questions regarding their professional and educational background. In the second part, they are presented with 4 scenarios, each containing 12 actions. They are supposed to choose which actions they would be more likely to perform in that specific scenario. They choose on a scale of 6 and there is an additional option named "I do not know". Finally, they are given a questionnaire to express their beliefs about classroom assessment and professional development. The instrument's reliability and validity have been checked by expert panels and various tests throughout the instrument development process. The instrument has been used in many different studies since its development to examine how teachers' assessment approaches are shaped under different variables such as career stages, subject area and professional development (DeLuca et al., 2016).

Finally, Kremmel and Harding (2020) developed a framework to meet the demands of the field in terms of empirical research. They pointed out that established models "remain theoretical in nature and represent the perspectives of language assessment researchers rather than stakeholders themselves," and that they have mostly been "prescribed sets of components based on principles of good practice." (Kremmel &

Harding, 2020, p. 100). They attempted to fill this gap by designing the “Language Assessment Literacy Survey”, which they later administered to 1086 participants from 77 countries. The model revealed many other dimensions of assessment literacy and emphasized that LAL is “multidimensional and developmental” and that “different levels of expertise or specialization will require different levels of (language assessment) literacy” (Pill & Harding, 2013, as cited in Kremmel & Harding, 2020, p. 102). The survey also identified a variety of roles, such as teachers, test developers, examiners, researchers, policy makers, and parents. With this model, Kremmel and Harding (2020) managed to provide a more inclusive measurement of teacher assessment literacy and stepped out of the teacher centered frameworks.

For the purposes of this study, the ACAI model was used to examine how in-service EFL teachers approach classroom assessment. It was especially suitable for examining how teachers understand and apply classroom assessment in their context thanks to its multifaceted structure.

2.6. Studies Conducted on TAL and LAL

Since AL has been widely recognized, empirical studies play a key role in presenting a broader understanding of how teachers understand and implement assessment practices. Many studies have been conducted to examine teachers’ knowledge, beliefs and practices in diverse educational settings and find the impact of these factors on their assessment literacy. These studies help the stakeholders understand teachers’ needs and the assessment challenges that they face in their everyday practices. This section reviews key empirical studies that measure teachers’ assessment literacy. The studies are divided into two sections: studies conducted internationally, and studies conducted in Türkiye. This division aims to present global trends and locally identified factors that shape assessment literacy.

2.6.1. International Studies Conducted on TAL and LAL

One of the earliest studies conducted on teachers’ assessment literacy was by Plake, Impara and Fager in 1993. The purpose of the study was to evaluate teachers’

assessment competencies based on the seven standards established by the Standards for Teacher Competence in Educational Assessment of Students (AFT et al., 1990) and to investigate how prior training in assessment affected their competency levels. They used a specially designed instrument for their study. There were two parts in the instrument. The first part consisted of 35 items designed to measure teachers' knowledge of assessment, and the second part consisted of some professional background questions and questions related to their views of assessment. 555 teachers from all around the USA took part in the study. In the first part, the teachers scored high in applying, grading and analyzing test results. However, they scored the lowest in communicating assessment outcomes. In addition, they struggled in "using assessments in grading" and "increasing the reliability score of a test" (Plake et al., 1993, p. 11). In the second part, the majority of the teachers agreed that teacher-developed instruments should be used to enhance learning. Overall, the teachers who had prior training on assessment scored higher than the ones who did not have prior training. They emphasized the need for sustainable and contemporary training and the lack of adequate preservice teacher training. Plake et al. (1993) noted that further training was needed to fully understand the needs of the teachers to improve professional development and classroom practices and contextual factors should be investigated to examine their effects on TAL.

Mertler (2003) conducted a study to examine the assessment literacy of preservice and in-service teachers by building on Plake et al. (1993). He created the Classroom Assessment Literacy Inventory (CALI), which was also a knowledge-based tool adapted from Plake et al.'s instrument. The first part consisting of 35 knowledge-based items was slightly revised by changing the wording in some items and 7 demographic items were added. The participants were 67 preservice and 197 in-service teachers. Preservice teachers performed best in selecting suitable assessment methods, and in-service teachers performed best in applying, grading and analyzing assessment outcomes. Both groups performed worst in "developing valid grading procedures" (Mertler, 2003, p. 18). In-service teachers proved to be more assessment literate in almost all of the standards. The findings showed that there was an inevitable need for targeted training in valid grading procedures, and that experience contributed positively to assessment literacy. In addition, the results raised doubts

about the efficiency and adequacy of teacher education programs in terms of preparing pre-service teachers for the requirements of the job regarding assessment. Mertler (2003) suggested that teachers' professional development needs needed to be explored further to develop effective training programs.

Ashraf and Zolfaghari (2015) conducted a study to examine if and how assessment literacy was related to Iranian teachers' age and teaching experience. The data collection tool was a 50-item self-reported questionnaire developed based on Michigan's Assessment Literacy Standards (2013). 658 Iranian teachers completed the questionnaire. The findings revealed that assessment literacy increases with both age and experience, and a significant increase was observed in teaching practices and knowledge. Also, the results supported the claim that assessment literacy evolved through practice over time. Ashraf and Zolfaghari suggested that PD programs should be designed by keeping the impact of teaching experience and age in mind.

Jannati (2015) carried out a research study to investigate the perception and practice gap in the field. She decided to examine the effect of teaching experience on perceptions and use of assessment. A researcher-developed semi-structured interview was used as the primary data collection tool. 18 Iranian EFL teachers from diverse teaching contexts participated in the study. The participants stated that assessment was an essential part of language education programs. They said that since assessment affects students' motivation, they should be knowledgeable about the assessment procedures in order to prevent the prominent backwash effect. They also stated that even though contextual factors such as student population and institutional constraints affected their use of formative and summative assessments from time to time, they used a mix of them. Their teaching experience did not play a significant role in their self-reported perceptions or practices of assessment. Jannati (2015) suggested that PD programs should be designed to bridge this gap between theory and practice, and more research should be done to investigate it further.

Similar to Jannati (2015), Xu and Brown (2017) carried out a study to investigate how factors such as age, experience and training affected AL. They used an adapted version of the Teacher Assessment Literacy Questionnaire (Plake et al., 1993), and

they worked with Chinese English teachers at the higher education level. 891 English teachers from different institutions took part in the study. The majority of the participants displayed a basic competency level in various domains such as “aligning tasks to goals, grading, interpreting standard scores” (Xu & Brown, 2017). No significant main effects have been identified based on age, experience, gender or training. The study highlighted the need for more contemporary AL standards in China and continuous and sustainable PD activities (Xu & Brown, 2017).

Berry et al. (2019) contributed to the AL research by conducting a study to understand what teachers did in the classroom, how they described their practices and what kind of support they needed. The study was conducted with 54 EFL teachers from the UK, France and Spain. The data collection process was completed in three stages. In the first stage, they carried out semi-structured interviews with EFL teachers. In the second stage, they did classroom observations and follow-up interviews with some teachers. In the final stage, they created focus groups with teachers from a global language teaching institution. They organized their findings according to Davies’ framework (Davies, 2008). The findings revealed that teachers understood the importance of assessment, but they did not like the connotations of testing. They felt confident in teaching, but they lacked confidence in assessment. Furthermore, even though they used many alternative assessment methods, they did not realize that those methods were also considered as assessment. Most of the participants labeled those practices as “good teaching” (Berry et al., 2019, p. 118). Therefore, Berry et al. (2019) suggested PD activities that would help teachers understand the terminology and reframe what they already knew and practiced.

Sultana (2019) examined the LAL in a highly exam-driven context, Bangladesh. 10 English teachers from different schools took part in the study. She aimed to understand how English teachers understand and apply assessment and if they were “academically and professionally ready” to carry out assessment-related tasks and responsibilities (Sultana, 2019, p. 1). The participants answered some questions related to their academic background, perceptions and use of assessment, and then they participated in a semi-structured interview. The results showed that teachers had little to no training in assessment. Thus, they developed their own understanding of

assessment over time with the help of experience, advice from colleagues and sample exam papers. They did not use any alternative assessments, and all of their practices were in line with the exam in their context. Furthermore, they perceived assessment as a grading tool and indicated that they were happy with their existing practices and did not need any additional LAL training. According to Sultana (2019), serious efforts should be made to provide assessment training to EFL teachers in Bangladesh to address this problem.

Kremmel and Harding (2020) developed the “Language Assessment Literacy Survey”, which they later administered to 1086 participants from 77 countries. The model revealed many other dimensions of assessment literacy and emphasized that LAL is “multidimensional and developmental” and that “different levels of expertise or specialization will require different levels of (language assessment) literacy” (Pill & Harding, 2013, as cited in Kremmel & Harding, 2020, p. 102). These dimensions were from Taylor’s (2013) original framework, but they also expanded it. The findings revealed that other stakeholders reported higher needs compared to teachers. With this model, Kremmel and Harding (2020) managed to provide a more extensive measurement of teacher assessment literacy, which includes all stakeholders.

Chen (2021) aimed to explore secondary school EFL teachers’ perceptions of assessment, their LAL and training needs. Since it was a case study, she had a small sample size. Therefore, she used multiple data collection tools such as “focus group interviews, individual interviews, observations, lesson plans, journals and school documents” (Chen, 2021, p. 1). The findings revealed that teachers had diverse perceptions of assessment, and assessment practices showed variation across participants. The participants used formative assessment tools, but they did not directly link them to student performances. Also, it was reported by the participants that there was a lack of formal AL training in secondary schools.

Latif and Wasim (2024) investigated the classroom assessment practices of in-service EFL teachers. 90 teachers participated in their study, and the data were collected through classroom observations, document analysis and thematic analysis. The findings revealed that while the instructors emphasized the importance of formative

assessment practices and reported using various assessment tools. However, classroom observations showed that they relied on traditional assessment practices, and their use of formative assessment was limited due to institutional factors and the exam driven context. Therefore, their practices and beliefs did not align with each other, which showed that there was a clear gap between assessment literacy and professional development. Latif and Wasim (2024) recommended that teacher training regarding assessment should be strengthened, and assessment policies should be reformed.

Amirian (2025) investigated the Iranian teachers' formative assessment literacy (FAL). According to the findings, the teachers felt most confident in applying formative assessment tools. Teaching experience appeared to increase teachers' FAL. However, Amirian (2025) added that ethical and personal factors also contribute to the enhancement of FAL. Therefore, he recommended that teacher education programs should be designed not only to increase assessment knowledge by keeping the effect of teaching experience in mind but also to provide practical applications of formative assessment tools and to focus on issues related to ethics and personal assessment philosophy.

Several studies have also employed the ACAI framework to investigate teachers' assessment approaches and practices, and they can provide valuable points of comparison for the present study.

Coombs (2017) conducted a study in which he used an adapted version of the Approaches to Classroom Assessment Inventory (ACAI) (DeLuca et al., 2016) in order to explore how teacher trainers developed their assessment approaches across ACAI's four domains: purpose, process, fairness, and theory (DeLuca et al., 2016). 108 teacher trainers from different institutions participated in the study. The findings revealed that participants showed a high level of support for the assessment purpose and process themes and the lowest level of support for the assessment theory theme. The results showed consistency within themes, but it varied within groups. For instance, the teacher trainers who taught assessment courses showed a more balanced variation of support for the approaches in the assessment purpose theme compared to

trainers who taught professional studies. Also, the findings indicated that teacher education programs and educators should follow a more balanced approach to improve the coverage of assessment topics and broader follow up studies should be conducted to map trends and patterns in teacher trainers' assessment literacy (Coombs, 2017).

Coombs et al. (2018) investigated how teachers' assessment approaches change over four different career stages. They defined these career stages as “initial pre-service, beginning in-service, early in-service, and established in-service teachers” (p. 137). Coombs et al. (2018) worked with 727 teachers from Canada and the USA, and they divided them into balanced career groups. An adapted version of the ACAI was used as the primary data collection tool. The findings showed that assessment for learning, differentiated fairness and balanced theory were supported by teachers in all career groups. It was established that career stages significantly impacted all components of the ACAI framework except for assessment process. Furthermore, the teachers who were in the early years of their careers preferred summative assessments more than the teachers who were in the late years of their careers. The study had strong implications for AL research in the field, such as tailoring PD according to career stages and recognizing the importance of experience and context in understanding AL.

DeLuca et al. (2018) conducted another study with ACAI to examine “how teachers approach classroom assessment and how their perceived skill levels and professional learning preferences vary according to career stage, teaching division, and assessment education background” (p. 356). Also, they aimed to validate the ACAI framework, which was based on Classroom Assessment Standards (JCSEE, 2015). They worked with 404 teachers from Canada and the USA. The ACAI had 3 parts. The first one was scenario-based items that reflected the participants' assessment approaches. The second part was about their self-reported skills in terms of designing, implementing and interpreting assessments. The final part was about PD needs and learning preferences. The findings in the first part indicated that the majority of teachers preferred formative assessment over summative assessment. The majority of the teachers preferred scoring over assessment design and

communication in the assessment process theme. Similarly, most of them preferred equitable fairness in the assessment fairness theme. However, experienced teachers reported that they were more likely to perform the differentiated fairness approach. In terms of assessment theory, the majority of the teachers emphasized the validity approach. The findings in the second part indicated that the participants found themselves more competent in giving feedback and guiding instruction with the help of assessment. Yet, they found themselves less competent in communicating with parents and in classroom observations. The findings in the last part indicated that the participants wanted to learn more about fair assessment for diverse learners, and they would prefer collaborative and personalized PD activities. Coombs et al. (2018) stated that even though the study was limited to North American contexts, it was still valuable in validating the instrument.

DeLuca et al. (2021) conducted another study with ACAI to understand how cultural patterns and differences are understood and applied in assessment. They designed a cross-national survey with 710 teachers from the USA, Canada and China. They analyzed the results in relation to 4 domains of the ACAI framework. The findings showed that the approaches that received the lowest support from the participants were the use/scoring and standard fairness approaches. On the other hand, the approaches that received the highest support from the participants were AfL and design approaches. Also, Chinese and Canadian teachers showed lower levels of support for AoL compared to US teachers. Due to the high-stakes exams, US teachers tended to use a mix of formative and summative approaches dominantly. The results obtained from the study highlighted how cultural, curricular, and policy-related factors shape teachers' assessment approaches and advised localized professional learning and avoiding one-size-fits-all LAL methods (DeLuca et al., 2021)

Nayagi and Rajendran (2020) conducted their study in the Indian context to examine preservice teachers' approaches to classroom assessment. They used an adapted version of the ACAI (DeLuca et al., 2016) as their primary data collection tool, but they did not use part 3 of the questionnaire, which was about professional development needs and priorities. 131 undergraduate students participated in their

study. The findings indicated high levels of support for assessment for learning, assessment of learning and reliability. They also indicated low levels of support for design, validity and differentiated fairness. Gender affected the results as there was a notable difference between gender groups in assessment purpose and assessment theory themes. Furthermore, in the second part of the ACAI, they indicated having strong confidence in giving meaningful feedback, interpreting assessment results and promoting student autonomy in terms of self-monitoring. Nayagi and Rajendran (2020) stated that even though the assessment courses increased the awareness and knowledge of assessment, practical training was definitely needed to improve students' practical skills.

Another study that used ACAI was conducted by Lazarakou (2025) in Greece. The study aimed to investigate Greek pre-service teachers' approaches to classroom assessment. 515 pre-service teachers took part in the study. Only Part A of the ACAI, which is related to how teachers view and approach assessment, was utilized. The findings showed that preservice teachers viewed assessment as a tool for learning and for improving instruction, and they thought that promoting self-monitoring in students was important. Also, they expressed hesitation towards large-scale standardized testing, and they were concerned about the anxiety that was caused by assessment (Lazarakou, 2025, p. 86). Based on the findings, Lazarakou suggested that teacher education programs should support more student-centered, differentiated and contextualized assessment strategies and the assessment policies should be adapted to reduce assessment-related stress and promote a more positive emotional climate in the classroom.

In summary, the studies demonstrate that teacher assessment literacy is influenced by various factors such as educational background, teaching experience and cultural, contextual and institutional factors. Studies also show that teachers are more competent in some areas of assessment and need support in others. Many teachers' ability to apply, analyze and interpret assessment is stronger than their ability to communicate assessment results to other stakeholders and to use formative assessment to improve students' performance. Furthermore, some studies conducted with pre-service teachers reveal that teacher candidates lack knowledge regarding

practical applications of assessment, and they do not receive sufficient training in pre-service education programs. Findings show that teaching experience may have an impact on TAL to some extent as it generally enhances assessment literacy. However, its effects are not always consistent, and it is highly dependent on the educational background and context. Therefore, studies point out that there is an undeniable gap between theoretical knowledge and practical applications, and teachers express a desire for practical, context-oriented and collaborative PD activities and programs.

2.6.2. Studies Conducted on TAL and LAL in Türkiye

This subsection presents some key studies conducted on TAL and LAL in Türkiye.

Hatipoğlu (2015) conducted a study to investigate the assessment training preservice English teachers were provided with, their expectations from the English Language Testing and Evaluation courses and their self-reported needs regarding assessment. 124 preservice English teachers took part in the study. The participants were given a survey, and they were asked to participate in a semi-structured interview. The findings revealed that preservice teachers received very limited training considering that they only took one course related to assessment during their preservice teacher education. Furthermore, many of the students could not identify any areas to work further on regarding assessment, and they lacked awareness of assessment methods. Their understanding of assessment was mainly constructed by the examination centered nature of Türkiye's educational system. However, despite all of these, they were aware of the importance of assessment training for their future teaching roles. Based on the findings of the study, Hatipoğlu (2015) suggested that more courses focusing on assessment should be integrated into the preservice teacher education curriculum, and pre-service teachers should be provided with more hands-on and collaborative training.

Kavaklı and Arslan (2019) conducted another study to explore the assessment literacy levels and perceptions of pre-service teachers. The study was conducted with 36 4th-year students in an English Language Teaching department at a university in Türkiye. Assessment Literacy Survey (Volante & Fazio, 2007) and European

Network of Language Testing and Assessment (ENLTA) student questionnaire were used to collect data. The findings showed that the participants reported they had a moderate level of AL, and most of them viewed assessment in a traditional way. Also, they expressed their desire for a more “active role” in the assessment processes. Therefore, they stated that they needed training during which they could put the theoretical knowledge into use. Based on the results, Kavaklı and Arslan (2019) pointed out that ELT programs should provide the students with practice-oriented courses that mirror real-life assessment development. Moreover, the students should be introduced to alternative assessment approaches so that their understanding of assessment wouldn't be limited to summative practices.

In another study that was conducted with in-service EFL teachers, Öz and Atay (2017) aimed to explore how Turkish EFL teachers view classroom assessment and reflect their AL into practice. Also, the study aimed to examine the effect of experience on the participants' views of assessment. A semi-structured interview was used to collect data, and 12 in-service EFL teachers from a Turkish university participated in the interviews. Öz and Atay (2017) stated that all teachers understood the importance of classroom assessment, but they had differing perceptions and definitions. The majority of the teachers preferred formative assessment tools and methods, but they could not use it as frequently as they wanted to due to contextual factors. Also, there was a gap between their AL and classroom practices. They pointed out that teacher education programs should equip teachers with formative assessment tools that they can adapt and utilize in their teaching contexts, and institutions should provide meaningful training in assessment design, feedback and the use of rubrics (Öz & Atay, 2017).

Ölmezer-Öztürk's (2021) research study also focused on preservice EFL teachers. A focus group interview and two reflection reports were used to collect data, and 48 junior students, who were taking their first course of assessment, took part in the study. The participants stated that initially, assessment had negative connotations for them. However, this situation changed by the end of the course. They viewed assessment as a tool for grading at the beginning, but their perception changed over time. They started to view assessment as a tool to enhance student learning. The

course helped them familiarize themselves with alternative assessment methods. This shift proved the fact that hands-on and carefully constructed assessment training before graduation could make a difference in preservice teachers' LAL. Therefore, Ölmezer-Öztürk (2021) highlighted that undergraduate ELT education should have courses that combine theoretical base with practical applications.

Şahin and Hatipoğlu (2023) conducted a study to examine the current testing courses in the English Language Teacher Education Programs. They aimed to investigate the educational background and qualifications of the instructors, the materials that they used and the content covered in these courses. 21 instructors teaching testing courses participated in their study. They collected their data through semi-structured interviews, testing course syllabi and information given related to the testing courses on universities' websites. Based on the results, they found out that the majority of the instructors did not specialize in language testing and assessment, the topics covered in the course were not up to date and summative assessment dominated the course content. Therefore, Şahin and Hatipoğlu (2023) commented that the existing testing courses did not adequately prepare pre-service teachers for the classroom assessment procedures. They also emphasized that formative assessment should be studied in detail in testing courses, and instructors who specialize in testing and assessment should teach these courses.

In addition to the studies conducted with preservice English teachers, Büyükkarcı (2016) conducted a study with in-service English teachers to investigate their AL levels and to explore whether any factors such as experience or additional training affected their AL. They used the Assessment Literacy Inventory (ALI) (Mertler & Campbell, 2005) as their primary data collection tool. 32 in-service teachers took part in the study from primary schools and universities. The findings showed that ELT teachers had very low levels of AL, and they were not provided with sufficient training by the existing PD programs. As for the impact of teaching experience and postgraduate education, they did not lead to a significant change in the AL levels of the teachers. Büyükkarcı (2016) expressed that an improvement in both preservice and in-service teacher education programs was urgently needed, and the teachers should be provided with hands-on and long-term assessment training.

Yastıbaş and Takkaç (2018) conducted a study to investigate how EFL instructors design appropriate assessments serving different instructional goals. The participants were 8 EFL instructors working at a university. They used think-aloud protocols to collect data (p. 180). The findings indicated that the instructors were very coursebook dependent while preparing summative assessments, and they mainly focused on “content validity” (Yastıbaş & Takkaç, 2018, p. 188). Also, most of the participants indicated that they were concerned about the washback effect of their assessment practices and put their students’ needs first while designing assessments. They analyzed their exams with a critical lens and developed them accordingly. Based on the findings, Yastıbaş and Takkaç (2018) suggested that textbook reliance should be decreased, and assessments should be aligned with exam design related frameworks.

Çelebi and Kuşuçuran (2019) conducted another study with in-service EFL teachers and elementary school teachers to investigate their preferences for alternative assessment methods across some variables such as gender, their undergraduate education and the effect of in-service assessment training. They found out that English teachers were more aware of the alternative assessment methods, and they used them more frequently. Also, teachers showed varying preferences for alternative assessment tools based on the faculty that they graduated from. Education faculty graduates showed higher preferences for role-play, portfolio and oral presentation. On the other hand, science and literature department graduates showed higher preferences for multiple choice and true/false tests and peer assessment. In addition, since the science and literature department graduates did not take any courses on assessment during their undergraduate education, they showed lower preferences for gamification, discussion and the think-aloud method. Therefore, it can be said that undergraduate education affected the participants’ preference for certain assessment methods and tools.

Önalın and Gürsoy (2020) conducted a study to investigate how Turkish primary and secondary school EFL teachers implement assessments and interpret their results. 348 teachers from private and public schools took part in the study. A questionnaire developed by the researcher was administered to the participants. The results showed that teachers utilized assessment mainly for feedback purposes. With the help of the

assessment, they gave feedback to themselves about their own teaching, to the students and parents about achievements or improvement areas and to the administrators about the effectiveness of the program. The study also revealed a significant difference between experience groups. The group with the highest level of experience was reported to use assessment for the actual goals of assessment practices. The student involvement in assessment processes was not as much as it was desired to be, but the group with more experience demonstrated more positive attitudes towards it. Furthermore, private school teachers and experienced teachers reported using both summative and formative assessment more than public school teachers and teachers at the early stages of their careers. Önalın and Gürsoy (2020) suggested that providing public school teachers and teachers with little to no experience with more support and educational opportunities and organizing PD activities that were focused on student-centered assessment practices and interpreting assessment results to guide teaching could be beneficial to increase teachers' AL levels.

Işık and Sarı (2021) aimed to investigate the LAL levels of ELT teachers in Türkiye regarding their "pre-service education backgrounds (ELT, ELL, or ETI)" (p. 907). 180 ELT teachers participated in their study. They utilized an adapted version of Vogt and Tsagari's (2014) questionnaire. Their results indicated a low level of LAL and no apparent desire for assessment training among ELT teachers. The participants noted that since they did not have any assessment training, their experiences as educators and students affected how they practice assessment. They also reported that they lacked experience in testing listening and speaking and test preparation because they use ready-made exams or find the questions online. Finally, they expressed that they viewed assessment as more as a tool for grading rather than a tool for learning. Based on the findings, Işık and Sarı (2021) emphasized that raising awareness of inadequacies should be the first step in helping teachers increase their LAL levels, and practical and classroom-based PD programs should be implemented to improve teachers' assessment practices.

Kaya and Mede (2021) conducted a study to investigate the EFL instructors' knowledge regarding language assessment. 195 EFL instructors participated in their

study, and they used the Language Assessment Knowledge Scale (LAKS) to collect data. The findings showed that the instructors demonstrated high level of assessment knowledge. They showed the highest performance in reading, and speaking appeared to be the most difficult skill to evaluate. Also, the participants stated that they were aware of key components of assessment such as validity and reliability and they used formative, summative and alternative assessments. Kaya and Mede (2021) recommended that initial teacher training should be improved, and in-service assessment training should be reformed to equip teachers with the skills needed for classroom assessment practices. Also, they mentioned some challenges that they faced while designing and implementing assessments such as lack of student motivation, time constraints, lack of technological resources and parent expectations.

Çimen (2022) conducted a study with in-service EFL teachers to investigate how the teachers evaluate their students' performances, whether their assessment practices align with the curriculum requirements and how they feel about assessment. They collected data through semi-structured interviews, open-ended questions and sample pen-paper exams graded by the teachers. According to the findings, teachers mainly used assessment to track student learning. They claimed to assess all four components of the language in an integrated way, but the sample exam papers showed that they mainly focused on grammar, vocabulary and isolated skills. Also, most of the teachers used traditional pen-paper exams and they were the main feedback providers. Therefore, Çimen (2022) commented that the teachers' practices mainly did not match with the curriculum, and they needed assessment training in authentic and alternative assessment methods and multiple feedback sources.

Tunçer and Merç (2023) aimed to identify the possible gap between Turkish EFL teachers' self-reported and actual assessment practices in their study. Their findings indicated that the teachers claimed that they used formative assessment tools frequently in their teaching practices. However, the task-based questionnaire and interviews revealed some inconsistencies. According to the task-based questionnaire and interviews, the classroom discussions and the use of self and peer assessment were quite limited, and the feedback aimed to praise the students rather than to improve their performance. Tunçer and Merç (2023) commented that Turkish EFL

teachers were knowledgeable about formative assessment tools and methods, but their practices were limited. Therefore, there was an urgent need for in-service assessment training to provide the teachers with practical implementations of formative assessments.

Doğan and Ünal (2024) aimed to examine the LAL levels of EFL teachers working at K12 schools in Türkiye and analyze their needs. 272 EFL teachers took part in the study and the Language Assessment Knowledge Scale (LAKS), and semi-structured interviews were utilized for data collection. LAKS was designed to measure knowledge of various language skills such as reading and listening, but it also includes items related to assessment tools and methods. The interviews were conducted to find out whether there were any gaps between the knowledge and competence that teachers reported and their actual practices. According to the findings, the participants were skilled in testing reading, but they struggled in testing listening even though they viewed themselves as competent in that area. Therefore, it can be said that the competence that teachers reported and their actual competence levels do not always match. The lack of competence in listening was considered to be due to limited training in designing tests in listening. The teachers stated that they wanted collaborative and hands-on training and additional support for formative assessment.

Sarıyıldız Canlı and Altay (2023) carried out a study to investigate the training needs of in-service EFL teachers regarding testing and assessment. 300 participants took part in their study, and they completed an online questionnaire to indicate their training needs. The participants reported that they needed support in designing tests to measure students' achievements and proficiency levels, alternative assessments, summative and formative assessments, determining the achievement of the learning outcomes based on assessment and testing speaking and writing skills through various assessment tools. Therefore, Canlı and Altay (2023) commented that training across all aspects of assessment was needed, but the emphasis put on these aspects should be determined based on teachers' needs.

Fındıklı and Büyükkarcı (2024) aimed to investigate and compare the assessment knowledge that teachers in Türkiye and some European countries had. 94 teachers

participated in the interviews and filled out the Language Assessment Knowledge Scale (LAKS). Turkish teachers scored significantly lower than European teachers in general assessment knowledge levels. In terms of their educational background, European ELT teachers scored significantly higher than Turkish ELT teachers, and European teachers who took assessment training also outperformed their Turkish peers. However, no significant difference has been observed in terms of the assessment knowledge levels of ELT and non-ELT graduates. Fındıklı and Büyükkarcı (2024) commented that this might be due to the fact that Turkish teacher education programs could not sufficiently prepare teacher candidates regarding assessment and testing, and they should be improved to better equip teachers with practical knowledge and skills that they needed to increase their assessment knowledge.

Mutlu (2025) carried out a study to examine Turkish EFL teachers' perceptions and practices of assessment. 257 teachers took part in the study, and the data were collected through online questionnaires and semi-structured interviews. The findings revealed that teachers used assessment to track student progress, determine areas for improvement and give feedback. The teachers stated that they preferred alternative assessment tools even though formal exams dominated their practices. They emphasized the importance of giving meaningful feedback in a timely manner to improve their students' performances. Based on the results, Mutlu (2025) suggested that more comprehensive assessment training should be provided for in-service EFL teachers, and there should be more alignment between teachers' practices, curriculum and assessment principles.

Yeni-Palabıyık and Daloğlu (2025) conducted a study to investigate the alignment and inconsistency between the policy requirements and Turkish EFL teachers' assessment practices. 3 English teachers and 170 students participated in the study. The data were collected through classroom observation notes, semi-structured interviews and document analysis. The findings revealed that the policy required implementing a mix of formative and traditional assessments and using speaking and performance-based tasks to evaluate students' performance. However, the teachers' implementations were dominated by traditional exams, and the speaking and

performance tasks were based on memorization. Therefore, while the policy aimed to implement a formative assessment-oriented approach, the actual practices remained summative in nature. Yeni-Palabıyık and Dalođlu (2025) commented that teachers' perceptions of assessment might have played a role in this conflict, and they emphasized the importance of raising teachers' awareness regarding assessment practices and policy requirements, effective training and strong teacher support systems.

Overall, the studies conducted in Türkiye reveal that the assessment literacy levels of both pre-service and in-service teachers are generally low, and the existing teacher education programs do not provide sufficient training in assessment. Pre-service teachers generally take one course on assessment, and it does not effectively prepare them for classroom assessment. Therefore, the studies highlight that pre-service teachers should be provided with opportunities that allow them to turn theoretical knowledge into practical applications, and their awareness of alternative assessment methods should be increased. In the case of in-service teachers, the studies show that teaching experience does not affect their AL levels significantly. It was also reported that teachers wanted to use formative assessment tools and methods, but their use was limited due to institutional and contextual factors. These findings highlight the urgent need for practical, collaborative and context-oriented professional development.

CHAPTER 3

METHODOLOGY

This chapter presents research design, research questions, data collection tools and procedures, data analysis, trustworthiness and the role of the researcher.

3.1. Research design

This research study is an explanatory case study with an embedded mixed methods approach. It aims to investigate the assessment approaches of in-service EFL teachers working at a private school in Ankara, to examine how teaching experience affects their assessment orientations and how EFL teachers perceive classroom assessment and reflect on their practices. Creswell and Creswell (2018) define case study as a design of inquiry through which a researcher explores “a real-life contemporary bound system” in detail using a variety of information sources. The case in this study is bounded by its setting, which is a private middle school in Ankara, and by the group of in-service EFL teachers who work there. This specific research setting was selected because it provides a controlled environment to explore assessment approaches of in-service middle school EFL teachers within the institutional culture. Additionally, since this study investigates the assessment approaches of in-service EFL teachers in one private institution, it is considered a single-site case study (Creswell & Creswell, 2018). Different categorizations have been made about the case study methods. However, when Yin’s (2018) categorization is taken into consideration, this study adopts the explanatory case study design. Explanatory case study approach is designed to analyze how and why relationships exist within real-world contexts (Yin, 2014). Since this study aims to understand why certain assessment orientations exist and how teaching experience affects teachers’ assessment orientations, explanatory case study design was considered appropriate to be used.

In addition, this study adopts an embedded mixed methods approach because the data were collected through the ACAI (DeLuca et al., 2016) questionnaire and semi-structured interviews. According to Creswell and Creswell (2018), mixed methods design offers a more comprehensive grasp of the research problem by combining both qualitative and quantitative data.

3.2. Research setting

The study was conducted at a private middle school in Ankara, Türkiye. The school was founded in the 1980s. The main campus is located in Ankara, but there are branch schools in 4 different cities. In addition to the branch schools, the institution provides educational guidance to several private schools around Türkiye. There are more than 900 teachers and over 8000 students on the Ankara campus and branch schools combined.

The study was conducted with English teachers in the middle school English department in the Ankara campus. There are 24 teachers in the department, including a department head/middle school English coordinator, a head of the academic development unit and an assistant principal. They do not teach due to their administrative duties.

The school has a total student population of approximately 1120, most of whom come from relatively high socio-economic backgrounds. Students' English proficiency levels vary by grade, and they range from A2 to B2 according to the Common European Framework of Reference (CEFR) scale. English is taught in all grade levels, and students have ten hours of English instruction per week. There are mixed-ability learners in classes, and the English curriculum is designed in alignment with CEFR guidelines. English is mainly used as the medium of instruction in all grade levels in the English lessons.

The school has a testing unit, consisting of teachers with testing training, that develops and oversees school-wide assessment policies. These policies provide clear guidance on both formative and summative assessment practices. Also, it provides

the coordinators, school administrators and teachers with an analysis of each summative and diagnostic assessment. In addition, there is another testing unit within the English department. This unit consists of the middle school department head/English coordinator and English teachers from the Ankara school and branch schools. The unit is responsible for preparing summative assessments that are implemented in all schools. All students on the main campus and in the branch schools take the same exams and quizzes unless they have an Individualized Education Plan (IEP). If they have an IEP, their exams are prepared by their own English teachers. Therefore, it can be said that the teachers in the English department do not have autonomy over designing summative assessments. However, a few teachers from each grade level in the Ankara school are invited to proofread and give feedback on summative assessments a few days prior to the assessment by the department head/middle school English coordinator. They have the right to ask for changes, adaptations and omissions of items in the assessments as long as they have valid reasons, and they evaluate the appropriateness and difficulty level of the reading texts, listening tracks and writing topics given in the exams and quizzes. The testing unit then revises and edits the items based on their feedback. The department head/middle school English coordinator tries to invite different teachers every time, but it may not always be possible due to the availability of the teachers during the day because of their busy schedules. The exam is not revealed to all of the teachers in the department due to exam safety procedures.

There are three standard procedures that are implemented in the department to ensure consistency and equality in scoring. Before each speaking and writing exam, three papers or audio recordings (one from a high-achieving student, one from an average/at-level student, and one from a struggling student) are shared with teachers across all schools. Teachers score these papers and recordings and send their grades to the department head/middle school coordinator. She takes a look at the grades, scores the sample papers/recordings and sends her own grades to all teachers. The teachers are then expected to look at the standardized scores and adjust their scores accordingly. For example, if a teacher awards extra points in the content category, s/he is supposed to review his/her scores and talk to the middle school English coordinator if any confusion arises. The second process is called blind double

grading, and it is only implemented in grading writing. The teachers score their own students' papers without writing the scores on the papers themselves and give the papers to another colleague from the same grade level that is randomly chosen by the department head/middle school coordinator. That colleague also scores those papers without writing the scores on the papers as well. Then, they come together and compare their scores. If the difference is greater than 2 points, the papers are scored by another person. The final procedure is spot-checking. After the scoring is finalized by the teachers, the department head/middle school English coordinator chooses some students from each class randomly and asks the teachers to give those students' papers/recordings to her. She scores the papers/recordings and checks the scores given by the teacher for any inconsistencies.

The formative assessments included in the lesson plans are entirely prepared by the teachers. Units and tasks are allocated to the teachers on the main campus and branch schools at the beginning of each academic year during the seminar period. The teachers who are responsible for preparing a certain unit or lesson plan prepare the formative assessment tools as well. They are encouraged and required to design a range of assessment tools such as exit tickets, peer and self-assessment, etc. Then, the department head/middle school English coordinator gives feedback to these plans and formative assessment tools, and they are revised accordingly. The plans and formative assessment tools are shared with the teachers so that they can all use them. If a teacher wants to redesign, adapt or change a formative assessment tool or a lesson plan, s/he can do it so long as it serves the same objectives and the tool or the plan is shared with all the teachers in the same grade level.

Professional development plays a key role in maintaining instructional quality. Teachers with institutional experience are required to complete at least three online training sessions, and new teachers are supposed to complete at least six training sessions that are selected by the department head/middle school English coordinator through an online platform. Also, their performance is evaluated through a school-wide system called the performance evaluation system which was developed based on The Framework for Teaching: Evaluation Instrument (Danielson, 2013). Teachers with institutional experience are observed once each term by the department head,

and they are required to do a peer observation once a year. New teachers are observed twice a term by the department head, multiple times by the head of the academic development unit and once a year by the school principal. They are also required to do peer observations and self-reflections by recording and watching their own lessons at least once a year. However, if needed, the head of the academic development unit may ask them to do multiple peer observations depending on their professional development needs. In addition, they have weekly meetings with the head of the academic development unit during which they talk about their teaching and assessment practices, their adaptation to the institution, their lesson plans, and their PD needs. All teachers are supposed to design their own plans and formative assessment tools for their observed lessons. They prepare them and have a pre-conference with the department head/middle school English coordinator during which they talk about their student profile, classroom routines, activities, and assessment tools. After the observation, both parties fill out a form that is developed based on the criteria indicated in *The Framework for Teaching: Evaluation Instrument* (Danielson, 2013) to evaluate the teacher's performance. Then, the observed teacher has a post-conference with the department head/middle school English coordinator to get detailed feedback regarding the lesson plans, activities and assessment practices. After the feedback session, the teachers set goals for themselves to work on the areas that need improvement.

This institution was chosen as the research site because it has a structured and well-established approach to assessment, and since it is a large private school, it has standardized practices across all campuses. It makes the institution a suitable research setting for an explanatory case study because it allows for an in-depth exploration of how teachers approach assessment and how teaching experience affects these practices in a clearly defined institutional framework.

3.3. Participants

As previously mentioned, the English department of the target institution consists of 24 teachers (including the researcher). 3 teachers were excluded from the sample because they did not teach due to their administrative duties, and the researcher did

not take part in the study either, so she was also excluded from the sample. Therefore, out of 20 teachers, 18 teachers were invited to participate in the study.

Purposive sampling was used to select participants for this research study. Purposive sampling is used to select participants that can provide useful information, and it is a technique that allows researchers to use resources effectively (Kelly, 2010; Palinkas et al., 2015). Criterion sampling was adopted among purposive sampling techniques (Campbell et al., 2020). The predetermined criteria were that the teachers had to have at least 5 years of teaching experience, and they had to work at the target institution for at least a year.

The participants were classified into two groups by following Hakim's (2015) classification: moderately experienced and highly experienced teachers. The first group consists of teachers with 5-10 years of teaching experience, and the second group consists of teachers with 11-15 years of teaching experience. There is no novice teachers group (0-4 years of teaching experience) because there are no novice teachers working in the English department of the target institution. Demographic information about the participants can be seen in Table 3.1 below. The table includes information about their age, gender, teaching experience, institutional experience and educational background.

Table 3.1. Demographic profile of the participants

Participant	Gender	Age	Teaching experience	Institutional experience	Education
P1	Male	31	9	3	BA and MA in English language and literature
P2	Female	32	8	1	BA in American Culture and Literature
P3	Male	29	6	4	BA and MA in English Language Teaching
P4	Female	33	10	1	BA and MA in Translation Studies
P5	Female	27	5	1	BA in English Language and Literature
P6	Female	30	7	3	BA in English Language and Literature
P7	Female	29	6	1	BA in Translation Studies

Table 3.1. (continued)

P8	Female	31	8	1	BA in English Language and Literature
P9	Female	27	6	6	BA in English Language Teaching
P10	Male	35	13	8	BA in English Language and Literature
P11	Female	39	14	2	BA in American Culture and Literature
P12	Female	30	14	3	BA in English Language Teaching
P13	Female	42	15	8	BA and MA in English Language and Literature
P14	Female	44	15	1	BA and MA in English Language and Literature
P15	Female	39	15	6	BA in English Language Teaching
P16	Female	47	14	2	BA in English Language and Literature
P17	Female	52	13	13	BA in English Language and Literature
P18	Female	39	12	1	BA in English Language Teaching

As it can be seen above, there are 3 male and 15 female participants. Their ages range from 27 to 52. Their teaching experience varies between 5 and 15 years. Half of the participants have 5-10 years of teaching experience, and the other half has 11-15 years of teaching experience. Their institutional experience ranges from 1 to 13 years with an average of approximately 4 years. Out of 18 participants, 9 have degrees in English Language and Literature, 5 in English Language Teaching, 2 in American Culture and Literature and 2 in Translation Studies. Among these participants, 3 have master's degrees in English Language and Literature, 1 in English Language Teaching and 1 in Translation Studies.

3.4. Data collection

Before collecting data, ethical approval was granted by the Human Subject Ethics Committee at Middle East Technical University with the protocol number 0327-ODTÜİAEK-2025 (Appendix A). As the researcher is also employed at the target institution, permission to conduct the study was obtained from the school administration prior to data collection.

3.4.1. Quantitative data collection instrument and data collection process

The Approaches to Classroom Assessment Inventory (ACAI) was utilized in the current study as the quantitative data collection tool. The latest version (v3.0) of the ACAI was used in the current study, and permission to use the ACAI was obtained from Dr. DeLuca via e-mail prior to data collection.

The ACAI framework was developed by Dr. Christopher DeLuca and his team at Queen’s Faculty of Education. In order to develop the instrument DeLuca et al. (2016) analyzed “15 assessment standards (1990–present) from six geographic regions (United States, Canada, United Kingdom, Europe, Australia, and New Zealand)” (p. 250). As a result of their analysis, they identified four themes that represent essential dimensions of TAL. These themes include assessment purpose, assessment process, assessment fairness and assessment theory. Then, each theme was associated with three priority approaches which were systematically determined as a result of their analysis of standards. Assessment literacy themes and approaches can be seen in Table 3.2 below:

Table 3.2. Assessment themes and approaches in ACAI

Theme	Approaches
Assessment purpose	Assessment of learning (AoL)
	Assessment for learning (AfL)
	Assessment as learning (AaL)
Assessment process	Design
	Use/scoring
	Communication
Assessment fairness	Standard
	Equitable
	Differentiated
Assessment theory	Consistent
	Contextual
	Balanced

Table adapted from DeLuca et al. (2016) - Approaches to Classroom Assessment Inventory: A New Instrument to Support Teacher Assessment Literacy.

As defined by DeLuca et al. (2016), the assessment purpose theme refers to selecting a suitable method of assessment to assess specific learning goals. There are three

priority approaches under this theme which are assessment of learning (AoL), assessment for learning (AfL) and assessment as learning (AaL). AoL approach refers to using evidence of learning to evaluate student learning and to score students' performance based on the accomplishment of the learning goal. AfL approach refers to using evidence of learning to inform instruction and identify learning gaps. AaL approach refers to promoting student autonomy in learning processes. The assessment process theme refers to the design, application and grading of assessments as well as the analysis and communication of results to stakeholders. There are three priority approaches within this theme which are design, scoring and communication. The design approach refers to the construction of reliable assessment tools to assess student performance. The scoring approach refers to focusing on adaptation and implementation of scoring procedures. The communication approach emphasizes the analysis and interpretation of assessment outcomes and feedback given to stakeholders based on these outcomes. The assessment fairness theme refers to ensuring that assessment is fair for all students. It consists of standard, equitable and differentiated treatment approaches. The standard treatment approach requires using the same criteria for all students. The equitable treatment approach advises adaptations or modifications to the assessment protocol for students with individualized education plans. The differentiated approach suggests tailoring assessment based on individualized needs and learning aims. The assessment theory theme refers to issues related to validity and reliability. It consists of consistent, contextual and balanced theory approaches. The consistent theory approach highlights consistency in grading, designing and implementing assessments. The contextual theory approach focuses on contextual factors while aligning assessments with the curriculum. The balanced theory approach combines these two approaches by focusing both on consistency and contextual factors (DeLuca et al., 2016).

The instrument consists of three parts. The first part includes items about the demographic and educational background of the participants. The participants are asked to provide information about their age, gender and country of residence. The questionnaire has two versions tailored for in-service and pre-service teachers. The pre-service teacher version focuses on the importance of teacher education while the

in-service teacher version focuses on current assessment knowledge and practice. The second part includes scenario-based items designed to examine the assessment approaches of the participants. In this part, the participants are given 4 scenarios representing “assessment dilemmas faced by teachers related to (a) summative assessment, (b) grading, (c) differentiated assessment, and (d) integrated assessment” (DeLuca et al., 2016, p. 251). The scenarios include situations where teachers are asked to make decisions about designing and applying summative assessments, deal with issues related to academic dishonesty, adapting the assessment to respond to diverse learner needs and planning an instructional unit that integrates assessment with teaching and learning. Each scenario includes 12 actions that align with the assessment themes and priority approaches explained above. The participants are asked to indicate the likelihood of performing each action listed under each scenario on a Likert scale from 1 – highly unlikely to 6 – highly likely and with an additional option titled “don’t know” (see Appendix B). The last part is about participants’ beliefs regarding classroom assessment. The items are aligned with the four themes of assessment literacy specified within the framework. It aims to conceptualize the underlying assumptions and attitudes that influence classroom assessment practices.

DeLuca et al. (2016) employed an expert panel method to establish the construct validity of the instrument. They invited “24 North American educational assessment experts”, and they made sure that “experts were recruited based on the following criteria: (a) membership in the Canadian Education Researchers’ Association or NCME; (b) published in the area of classroom or large-scale assessment; and/or (c) an active member of a university, school district, or education/assessment-related institution” (p. 256). Among these 24 experts, 10 of them agreed to participate. Also, a panel of educators consisting of 5 primary and 5 middle school teachers was formed. Each participant was asked to double rate each item in the scenario-based questions based on how well they aligned with the specified assessment literacy domain and priority approach. They used a scale from 1 – not aligned to 5 – strongly aligned. Then, the participants were asked to explain the reasons behind their ratings and make recommendations for the revision of the items if their ratings were ≤ 3 . For the parts about teachers’ confidence levels regarding classroom assessment and their professional development needs, the experts were asked to reflect on the

comprehensiveness, variety and complexity of the items and provide suggestions if necessary. When this process was finalized and the instrument was revised based on the feedback provided by the expert panel, the instrument was pilot tested with 404 teachers from Canada to provide evidence of validity and reliability of the instrument. The results revealed that there were no notable differences in response patterns between demographic groups in scenario-based items. Also, Parts 2 and 3 of the questionnaire showed high internal consistency with reliability coefficients that ranged from $\alpha = .74$ to $\alpha = .92$. Additional validity and reliability measures have not been conducted because the instrument had already undergone rigorous validity and reliability testing in its original development. The tool was adopted in the current research context, and the scenarios were used without modification to ensure consistency with the original framework.

The first part of the instrument was adapted to include items about institutional experience, educational background, and previous assessment training. The second part of the instrument, which includes scenario-based items, was used exactly as it was designed by Dr. DeLuca and his team. The researcher conducted piloting with 2 teachers who worked in the target institution but did not participate in the study and got their feedback. As a result, it was understood that no modifications were needed as they fit the current educational context. The last part of the study, which was about participants' beliefs regarding classroom assessment experiences, was not used in the study because the researcher chose to explore the perceptions and experiences of in-service EFL teachers through semi-structured interviews. Semi-structured interviews allowed for more flexibility in questions and an in-depth exploration of experiences and perceptions of the participants as they capture the rationale, contradictions and individual experiences of the participants, which aligned well with the qualitative nature of the third research question.

Since all participants were employed at the same school, the final version of the questionnaire was printed and given to participants face to face on the site. They were informed about the purpose of the research and assured about anonymity and confidentiality. Informed consent was obtained before the administration of the questionnaire. 18 questionnaires were collected to be analyzed.

3.4.2. Qualitative data collection instrument and data collection procedure

For qualitative data collection, semi-structured interview questions were developed to investigate in-service EFL teachers' perceptions and experiences of classroom assessment. The interview questions were adapted from Çayır (2022). She conducted her study with pre-service science teachers, so the items needed revisions and adaptations. The original questions focused on university education, anticipated practices and conceptual understanding of assessment within the context of science education. Since the current study focuses on in-service EFL teachers working at a private institution, the interview questions were revised and adapted to reflect their knowledge, professional responsibilities and contextual challenges.

The adapted version of the interview questions was developed under 7 categories: teacher background and experience, professional development, assessment practices and preferences, challenges in assessment, fairness in assessment, personal assessment philosophy and institutional context, and questions were changed, adapted, removed and added according to these categories. For example, the demographic questions were adapted to focus on teaching experience rather than academic background (E.g., "Which university do you attend?" was adapted to "How long have you been teaching English?"). The focus was shifted from science education to language teaching (E.g., "What do you think is the role and importance of assessment in science education?" was adapted to "What do you think is the main purpose of classroom assessment in ELT?"). University courses were replaced with professional development activities (E.g., "Have you taken any courses on assessment during your undergraduate education?" was adapted to "Have you taken any professional development courses on classroom assessment?"). Questions related to fairness, contextual factors and alternative assessments were added (E.g., "What does 'fair assessment' mean to you, and how do you ensure fairness?"). Finally, questions that required the participants to reflect on their understanding of assessment were added to gather deeper insights (E.g., "How would you describe your personal approach to assessment?"). These adaptations and modifications were made to keep the focus entirely on the ELT context and classroom assessment. Piloting was done with one teacher from the target institution's English department.

Based on the feedback that she gave, the term alternative assessment was clarified with examples in case the participants could not remember what they were, and an explanation regarding personal approach to assessment (principles that guide how a teacher designs, chooses, implements and analyzes an assessment) were provided to the participants in case they had difficulty in understanding the question.

17 participants took part in the semi-structured interviews. The interviews were conducted via Zoom or in person and in English or Turkish, depending on the participant's preference and lasted approximately 22 minutes. Interviews were recorded with participants' consent.

3.5. Data analysis

Quantitative and qualitative data were collected for the study. The quantitative data analysis included analyzing data that came from the ACAI questionnaire. Descriptive analysis (mean, standard deviation and frequency) was conducted through IBM SPSS Statistics 23.

Qualitative data from semi-structured interviews were analyzed using Braun and Clarke's (2006) inductive thematic analysis. According to Braun and Clarke (2006), thematic analysis includes six steps (p. 87):

1. Becoming closely acquainted with your data
2. Developing early-stage codes
3. Identifying themes
4. Refining themes
5. Creating titles and descriptions for themes
6. Presenting findings with the help of data

Following Braun and Clarke's (2006) framework, the interviews were transcribed verbatim. Then, the researcher familiarized herself with the data by reading each transcript multiple times and taking notes to determine repeated concepts or ideas. MAXQDA 2024 was used during the initial coding stage to develop phrase-level codes that reflect the meaningful parts of the participants' answers. Then, these codes

were reviewed, and emerging themes and subthemes were identified. The subthemes were grouped under suitable themes. The researcher kept a codebook (see Appendix F) throughout the process and updated it as the codes were named, grouped into themes and subthemes and defined. Themes were revised and refined until they clearly represented the repeated ideas and patterns in the data. Finally, excerpts were chosen based on how representative and detailed they were to represent key patterns in the data.

3.6 Trustworthiness

To ensure trustworthiness of the qualitative part of the research, 4 key elements established by Lincoln and Guba (1985), which are credibility, transferability, dependability and confirmability, were carefully considered.

Credibility refers to how well the findings reflect reality (Merriam, 2009). In this study, credibility was supported by including two data collection tools, namely the ACAI questionnaire and semi-structured interviews. Using both methods allowed the researcher to check how consistent the participants were in their responses. Since the researcher is also currently affiliated with the target institution, the potential researcher bias on data collection and interpretation were recognized. Therefore, member checking was conducted with all of the participants. After the interviews were transcribed and direct quotations were chosen, the transcripts and direct quotations were shared with the participants face to face, and they were asked to confirm the accuracy. All participants agreed that the transcripts were accurate, and the direct quotations represented their ideas.

Transferability refers to how generalizable the findings can be (Merriam, 2009). In order to support transferability, rich and thick descriptions of the research setting were provided so that the readers could determine the relevance of this study to their own contexts.

Dependability is similar to reliability and can be supported by audit trail, peer review and the researcher's role (Merriam, 2009). Peer examination was provided by the

thesis advisor at different times during the study. Although there is no formal audit trail, the research progress was clearly documented, and analyses were conducted systematically.

Confirmability refers to the extent to which interpretations can be supported with data (Merriam, 2009). It was supported by the inclusion of verbatim excerpts from semi-structured interviews.

In conclusion, even though not all strategies were applied, appropriate and transparent methods were used to ensure overall trustworthiness.

3.7. The researcher's position

Since a researcher's beliefs and backgrounds can affect the findings and interpretations of qualitative research, researchers are encouraged to openly share their beliefs, experiences, biases and backgrounds (Creswell, 2013). A researcher can be an outsider or insider in a research study depending on his/her position. I, as the researcher of this study, graduated from Middle East Technical University, and I have been working at the target institution for 6 years. Therefore, I would define myself as an insider.

In the first year of my career, I used to think that the most effective way to support learning and enhance student learning was to teach well. However, I have had the opportunity to observe many of my colleagues and have discussions about assessment and classroom activities with them during the time that I spent in this institution. Based on these observations and discussions, I realized that one of the most significant factors that contribute to student achievement and support learning was using assessment activities effectively, interpreting the results correctly and taking action based on the data provided by the assessment activities. This realization sparked my interest in assessment, and as of the first year of my teaching career, I have been working on my personal philosophy of assessment. Also, I have gained a deeper understanding of assessment with the help of the professional development activities provided by the institution and my studies throughout my master's degree

in English Language Education. Therefore, I wanted to explore what kinds of similarities and differences existed in the assessment approaches, experiences and perceptions of teachers working at the same institution. In addition, as a teacher who struggled with determining and clarifying her own approach to assessment in the early years of my career, I wanted to investigate whether teaching experience had any effects on this matter.

As an insider, I was able to establish a bond with the participants as I have been working with them for 6 years, and there are no power relationships between us. This helped me during the data collection process as the participants felt comfortable enough to share their experiences and thoughts with me openly. Also, they were able to share their experiences and perceptions with me openly thanks to my familiarity with the institutional assessment culture, assessment policy and everyday teaching and assessment practices. My understanding of shared practices helped me interpret the data more clearly. However, being an insider might also have had some potential disadvantages. The participants may not have shared some of their experiences and perceptions due to the bias of social desirability. Also, my shared understanding of experiences and practices could have led me to assume some common understandings or interpret the data through my own practices and beliefs. Since I was aware of this potential challenge, I tried to remain reflexive during the data collection and analysis processes, and I regularly reviewed the transcripts to avoid personal assumptions and to focus on the meaning intended by the participant. Finally, I used some strategies that were mentioned in the previous section to eliminate researcher biases and enhance the trustworthiness of my research.

CHAPTER 4

RESULTS

This chapter presents the findings collected from the questionnaires and the interviews conducted with in-service EFL teachers.

4.1. In-service EFL Teachers' Approaches to Assessment

The first research question, which was “*To what extent do in-service EFL teachers support a specific classroom assessment approach in relation to the four themes of assessment (assessment purpose, assessment process, assessment fairness, assessment theory) specified within the ACAI framework?*”, was answered by descriptive analysis of the questionnaire items. As previously mentioned in the methods section, following the ACAI framework and the six-point Likert scale used in the questionnaire, $M \geq 4$ was indicative of support for each approach. The M, SD, and percentages of the items were calculated and analyzed. Descriptive statistics for each scenario are given in Table 4.1 below.

Scenario 1: *You give your class a paper-pencil summative unit test with accommodation and modifications for identified learners. Ten of the 24 students fail.*

As a teacher in this situation, how likely are you to do each of the following actions as part of your overall response to the scenario?

The results from this scenario showed that participants demonstrated varying levels of support for the three approaches of the assessment purpose theme. They did not support a particular approach within this category in the specified scenario. The AoL action (“*keep and record the test grade for all students and move on to the next*

topic”) and the AfL action (“disregard the test grade, then reteach and retest the material”) both got a mean score of 2.78 (SD: 2.16, SD: 1.77). The AaL action (“have students generate a plan to relearn the material”) has a mean score of 3.39. Even though this approach received relatively higher support, there was not consensus among the participants, as the standard deviation (SD: 2.0) shows the participants had divergent views.

Table 4.1. Descriptive statistics for Scenario 1

Related theme	Approach	Actions	Percentage (%)							M	SD
			1 Highly unlikely	2	3	4	5	6 Highly likely	7 Do not know		
Assessment Purpose	AoL	Keep and record the test grade and move on to the next topic.	44.4	22.2	0	0	11.1	22.2	0	2.78	2.16
	AfL	Disregard the test grade, then reteach and retest the material.	38.9	11.1	11.1	16.7	16.7	5.6	0	2.78	1.77
Assessment Purpose	AaL	Have students generate a plan to relearn the material.	33.3	5.6	5.6	16.7	22.2	16.7	0	3.39	2.00
Assessment Process	Design	Analyze test results to determine if failure was the result of poor test design.	0	0	16.7	0	16.7	66.7	0	5.33	1.14
	Use/scoring	Remove test questions that most students struggled with and recalculate student scores without those questions.	55.6	11.1	11.1	11.1	5.6	0	5.6	1.83	1.38
	Commun.	Schedule student conferences to discuss test performance and next steps.	0	11.1	5.6	5.6	50	27.8	0	4.78	1.26
Assessment Fairness	Standard	Allow the entire class to retake a similar test.	33.3	22.2	5.6	11.1	22.2	0	5.6	2.50	1.72
	Equitable	Allow only students with exceptionalities who struggled with the test to take a new assessment.	33.3	27.8	22.2	5.6	11.1	0	0	2.33	1.33
	Differentiated	Provide a new assessment to every student who struggled with the test.	22.2	22.2	5.6	5.6	11.1	27.8	5.6	3.28	2.19
Assessment Theory	Consistent	Identify test questions that most students consistently answered incorrectly, then provide students with new questions to retest those concepts.	5.6	5.6	5.6	22.2	22.2	38.9	0	4.67	1.50
Assessment Theory	Contextual	Adjust student grades based on previously collected formative assessment information.	38.9	16.7	0	27.8	11.1	5.6	0	2.72	1.74
	Balanced	Analyze test results to identify test design errors and consider student circumstances for failure, then adjust individual student grades accordingly.	11.1	33.3	11.1	5.6	22.2	11.1	5.6	3.11	1.84

Participants’ responses to the assessment process theme indicated higher support for two of the approaches. The design action (“analyze test results to determine if failure was the result of poor test design”) got the highest level of support with a mean score of 5.33 (SD: 1.14). 83.4% of the participants stated that they were “likely” (16.7%) and “highly likely” (66.7%) to perform that action. It was followed by communication action (“schedule student conferences to discuss test performance and next steps”) with a mean score of 4.78 (SD: 1.26). The use/scoring action

(“remove test questions that most students struggled with and recalculate student scores without those questions got the lowest support”) (M: 1.83, SD: 1.38).

The assessment fairness theme also received low support from the participants in this particular scenario. The equitable fairness action (*“allow only students with exceptionalities who struggled with the test to take a new assessment”*) received the lowest level of support among the three approaches (M: 2.33, SD: 1.33). It was followed by the standard fairness action (*“allow the entire class to retake a similar test”*) with a mean score of 2.50 (SD: 1.72). The differentiated fairness action had the highest mean score with 3.28 (SD: 2.19). However, it does not show clear support for this action since the mean score is below the 4.0 threshold and 55% of the participants were unlikely to perform it. The responses to the assessment theory theme showed mixed levels of support. The participants supported the consistent theory action (*“identify test questions that most students consistently answered incorrectly, then provide students with new questions to retest those concepts”*) more than the other two approaches (M: 4.67, SD: 1.50). 61.1% of the participants stated that they were “likely” (22.2%) and “highly likely” (38.9%) to perform that action. It was followed by the balanced theory action (*“analyze test results to identify test design errors and consider student circumstances for failure, then adjust individual student grades accordingly”*) with a mean score of 3.11 (SD: 1.84). The contextual theory action (*“adjust student grades based on previously collected formative assessment information”*) got the lowest level of support with a mean score of 2.72 (SD: 1.74).

Scenario 2: *You discover that one of your students has plagiarized sections of their term project.*

As a teacher in this situation, how likely are you to do each of the following actions?

In the second scenario, all three of the actions in the assessment purpose theme received high support from the participants. The AaL action (*“ask them to reflect on why plagiarism is a problem and what they would do differently next time”*) received the highest level of support (M: 5.44, SD: 0.92). 83% of the participants stated that they were “likely” (16.7%) and “highly likely” (66.7%) to perform that action. It was

followed by the AfL action (“*have them re-write the plagiarized section in their own words, then re-grade the assignment*”) with a mean score of 4.89 (SD: 1.88). The AoL action (“*reduce the student’s grade*”) got the lowest level of support from the participants among these three approaches with a mean score of 4.89 (SD: 1.88).

Table 4.2. Descriptive Statistics for Scenario 2

Related theme	Approach	Actions	Percentage (%)							M	SD
			1 Highly unlikely	2	3	4	5	6 Highly likely	7 Do not know		
Assessment Purpose	AoL	Reduce the student’s grade.	5.6	5.6	5.6	16.7	27.8	33.3	5.6	4.39	1.82
	AfL	Have them re-write the plagiarized section in their own words, then re- grade the assignment.	5.6	5.6	0	0	27.8	50	5.6	4.89	1.88
	AaL	Ask them to reflect on why plagiarism is a problem and what they would do differently next time.	0	0	5.6	11.1	16.7	66.7	0	5.44	0.92
Assessment Process	Design	As the teacher, reflect on how the assignment could have been structured differently to deter plagiarism.	5.6	5.6	0	16.7	11.1	61.1	0	5.06	1.51
	Use/scoring	Adjust the student’s grade to reflect the portion of work that was plagiarized.	5.6	0	0	16.7	50	5.6	22.2	3.56	2.20
	Commun.	Discuss with the student the reasons for the plagiarism, severity of plagiarism, and negotiate potential next steps for their learning.	0	0	5.6	5.6	22.2	66.7	0	5.50	0.86
Assessment Fairness	Standard	Apply the same consequence you would for other students to ensure all students are treated equally.	0	0	0	5.6	22.2	72.2	0	5.67	0.59
	Equitable	Consider if the student has identified learning needs and exceptionalities before determining a response to plagiarism.	5.6	5.6	5.6	11.1	38.9	22.2	11.1	4.06	2.01
	Differentiated	Discuss why the student plagiarized and agree upon an appropriate alternative assignment.	11.1	16.7	11.1	5.6	44.4	11.1	0	3.89	1.68
Assessment Theory	Consistent	Apply all aspects of school policy on plagiarism to ensure consistency across all students.	0	0	5.6	11.1	16.7	61.1	5.6	5.11	1.57

Table 4.2. (continued)

Assessment Theory	Contextual	Consider the original aspects of the assignment and the plagiarized text to determine what the student knows and does not appear to know related to content expectations.	0	5.6	11.1	11.1	38.9	33.3	0	4.83	1.20
	Balanced	Apply the school policy on plagiarism unless there is a good reason that explains why the student plagiarized.	11.1	5.6	11.1	16.7	38.9	16.7	0	4.17	1.58

Participants’ responses to the assessment process theme reflected high support for assessment design and communication. Communication action (“*discuss with the student the reasons for the plagiarism, severity of plagiarism, and negotiate potential next steps for their learning*”) received the highest level of support (M: 5.50, SD: 0.86). 88.9% of the participants stated that they were “likely” (22.2%) and “highly likely” (66.7%) to perform that action. It was followed by the assessment design action (“*as the teacher, reflect on how the assessment could have been structured differently to deter plagiarism*”) with a mean score of 5.06 (SD: 1.51). The action that received the lowest level of support was the use/scoring action (“*adjust the student’s grade to reflect the portion of work that was plagiarized*”). It got a mean score of 3.56 (SD: 2.20).

In the assessment fairness theme, the standard fairness (M: 5.67, SD: 0.59) and equitable fairness (M: 4.06, SD: 2.01) approaches got high support from the participants. The standard assessment action (“*apply the same consequence you would for other students to ensure all students are treated equally*”) got the highest level of support. 94.4% of the participants were “likely” (22.2%) and “highly likely” (72.2%) to perform this action. It was followed by the equitable fairness approach (“*consider if the student has identified learning needs and exceptionalities before determining a response to plagiarism*”). However, while the mean score in this action was high, the standard deviation shows that the participants had divergent views. The differentiated fairness action (“*discuss why the student plagiarized and agree upon an appropriate alternative assignment*”) got the lowest level of support (M: 3.89, SD: 1.68).

The responses to the assessment theory theme also indicated high support for all three approaches. The action that received the highest level of support was the consistent theory action (“*apply all aspects of school policy on plagiarism to ensure consistency across all students*”). It got a mean score of 5.11 (SD: 1.57). 77.8% of the participants stated that they were “likely” (16.7%) and “highly likely” (61.1%) to perform that action. It was followed by the contextual theory action (“*consider the original aspects of the assignment and the plagiarized text to determine what the student knows and does not appear to know related to content expectations*”) with a mean score of 4.83 (SD: 1.20). The least supported action was the balanced theory action (“*apply the school policy on plagiarism unless there is a good reason that explains why the student plagiarized*”) with a mean score of 4.17 (SD: 1.58).

Scenario 3: *Out of 28 students in your class, 4 students are classified/identified with an exceptionality and have an Individualized Education Plan (IEP) (i.e., each student requires accommodations, but not a modified curriculum). Additionally, several other students have different learning needs that are not officially identified. You must decide how to accurately assess learning in your class.*

As a teacher in this situation, how likely are you to do each of the following actions?

Table 4.3. Descriptive Statistics for Scenario 3

Related theme	Approach	Actions	Percentage (%)							M	SD
			1 Highly unlikely	2	3	4	5	6 Highly likely	0 Do not know		
Assessment Purpose	AoL	Provide all identified students with accommodations on all summative assessments.	0	0	11.1	16.7	16.7	55.6	0	5.17	1.10
	AfL	Use formative assessments with all of your students based on their individual learning needs to support their learning.	0	0	11.1	5.6	38.9	44.4	0	5.17	0.99
	AaL	Allow each student to develop a learning plan based on their strengths and goals to support their own learning.	5.6	5.6	16.7	16.7	22.2	27.8	0	4.11	1.81

Table 4.3. (continued)

Assessment Process	Design	Design a variety of assessment methods to target each learning objective.	5.6	0	16.7	5.6	16.7	55.6	0	4.94	1.81
	Use/scoring	Develop and use different rubric criteria to assess identified students.	5.6	5.6	0	16.7	16.7	55.6	0	5.00	1.50
	Commun.	Explain to students and parents the purpose of accommodations and how they will be implemented to support student learning and assessment.	5.6	0	0	16.7	22.2	55.6	0	5.17	1.30
Assessment Fairness	Standard	Grade students using the same assessments and without accommodations.	66.7	5.6	11.1	0	5.6	11.1	0	2.06	1.80
	Equitable	Assume that accommodated assessments provide an equally valid measure of identified students' learning.	0	11.1	0	22.2	27.8	16.7	22.2	3.50	2.23
	Differentiated	Provide a variety of assessment options and allow students to choose how they will demonstrate their learning.	0	11.1	11.1	11.1	22.2	44.4	0	4.78	1.44
Assessment Theory	Consistent	Use the same scoring rubric for all students.	38.9	5.6	11.1	11.1	22.2	11.1	0	3.06	1.96
	Contextual	Consider individual student learning needs when assessing student work.	5.6	5.6	5.6	16.7	11.1	55.6	0	4.89	1.57
	Balanced	Use the same scoring rubric for all students but apply professional judgment to adapt criteria based on individual student needs.	5.6	0	27.8	11.1	33.3	22.2	0	4.33	1.41

In the third scenario, all three of the actions in the assessment purpose theme received high support from the participants. The AoL action (“*provide all identified students with accommodations on all summative assessments*”) and the AfL action (“*use formative assessments with all of your students based on their individual learning needs to support their learning*”) got the highest level of support. They both got a mean score of 5.17. These actions were followed by the AaL action (“*allow each student to develop a learning plan based on their strengths and goals to support their own learning*”), which got a mean score of 4.11 (SD: 1.81).

Participants' responses to the assessment process theme also indicated high support for all three approaches. The communication action (*“explain to students and parents the purpose of accommodations and how they will be implemented to support student learning and assessment”*) got the highest level of support from the participants (M: 5.17, SD: 1.30). 78.3% of the participants stated that they were “likely” (22.2%) and “highly likely” (55.6%) to perform this action. It was followed by the use/scoring action (*“develop and use different rubric criteria to assess identified students”*) with a mean score of 5.0 (SD: 1.50). Even though the design action (*“design a variety of assessment methods to target each learning objective”*) got the lowest mean score among all three (M: 4.94, SD: 1.81), it was still highly endorsed by the participants.

One of the actions received high support from the participants in the assessment fairness theme. The differentiated fairness action (*“provide a variety of assessment options and allow students to choose how they will demonstrate their learning”*) got the highest level of support with a mean score of 4.78 (SD: 1.44). 66.7% of the participants stated that they were “likely” (22.2%) and “highly likely” (44.4%) to perform this action. The other two approaches received low support, indicating weak endorsement of the actions stated within the related themes. The equitable fairness action (*“assume that accommodated assessments provide an equally valid measure of identified students’ learning”*) got a mean score of 3.5 (SD: 2.23), and the standard fairness action (*“grade students using the same assessments and without accommodations”*) got a mean score of 2.06 (SD: 1.80).

The responses to the assessment theory theme indicated high support for two of the actions and low support for one of them. The contextual theory action (*“consider individual student learning needs when assessing student work”*) received the highest level of support from the participants (M: 4.89, SD: 1.57). 66.7% of the participants stated that they were “likely” (11.1%) and “highly likely” (55.6%) to perform this action. It was followed by the balanced theory action (*“use the same scoring rubric for all students but apply professional judgment to adapt criteria based on individual student needs”*), which received a mean score of 4.33 (SD: 1.41). The consistent theory action (*“use the same scoring rubric for all students”*) got the lowest level of support among the three (M: 3.06, SD: 1.96).

Scenario 4: “You are preparing a new unit for your English class. As part of your planning, you must decide how to design assessments, how to use them during the unit, and how to evaluate student learning at the end.”

As a teacher in this situation, how likely are you to do each of the following actions?

Table 4.4. Descriptive Statistics for Scenario 4

Related theme	Approach	Actions	Percentage (%)							M	SD
			1 Highly unlikely	2	3	4	5	6 Highly likely	7 Do not know		
Assessment purpose	Assessment of learning	Start by designing a summative assessment based on curriculum expectations, then plan your lessons.	11.1	16.7	22.2	5.6	16.7	27.8	0	3.83	1.82
	Assessment for learning	Use formative assessments to guide your subsequent lesson planning.	0	5.6	0	16.7	27.8	50.0	0	5.17	1.10
	Assessment as learning	Co-construct learning goals with students, then have each student develop a personal learning plan for the unit.	11.1	27.8	11.1	38.9	5.6	5.6	0	3.17	1.38
Assessment process	Design	Design a summative unit assessment based on what was taught to students.	0	0	0	5.6	33.3	61.1	0	5.56	0.62
	Use/scoring	When planning a unit, consider when assessments and feedback will be provided to students.	0	0	0	0	27.8	72.2	0	5.72	0.46
	Communication	Discuss assignments and co-construct grading criteria for the unit with your students.	11.1	33.3	5.6	27.8	16.7	5.6	0	3.22	1.52
Assessment Fairness	Standard	Plan class lessons and assessments that are the same for all students and encompass the curriculum expectations.	11.1	16.7	11.1	16.7	16.7	22.2	5.6	3.61	1.94
	Equitable	Ensure that assessments throughout the unit are accommodated for formally identified learners	0	0	0	5.6	33.3	61.1	0	5.56	0.62
	Differentiated	Give all students a diagnostic assessment at the beginning of the unit and have students use their results to select personalized appropriate learning and assessment activities.	11.1	16.7	22.2	33.3	11.1	11.1	5.6	3.33	1.37

Table 4.4. (continued)

Related theme	Approach	Actions	Percentage (%)							M	SD
			1 Highly unlikely	2	3	4	5	6 Highly likely	7 Do not know		
Assessment Theory	Consistent	Work with other teachers to develop a reliable summative assessment for the unit.	0	0	0	16.7	22.2	61.1	0	5.44	0.78
	Contextual	In designing your assessments, prioritize the alignment between the assessment and the intended learning objectives.	0	0	0	0	27.8	72.2	0	5.72	0.46
Assessment theory	Balanced	Adapt assessments that have worked well with other students like yours to take into consideration the context and activities of your enacted lessons.	0	11.1	5.6	5.6	38.9	38.9	0	4.89	1.32

In the fourth scenario, one of the approaches in the assessment purpose theme was supported more by the participants. The AfL action (“*use formative assessments to guide your subsequent lesson planning*”) received the highest level of support from the participants (M: 5.17, SD: 1.10). 77.8% of the participants stated that they were “likely” (27.8%) and “highly likely” (50%) to perform that action. AoL and AaL received low support from the participants. The AoL action (“*start by designing a summative assessment based on curriculum expectations, then plan your lessons*”) got a mean score of 3.83 (SD: 1.82). The AaL action (“*Co-construct learning goals with students, then have each student develop a personal learning plan for the unit*”) got the lowest level of support with a mean score of 3.17 (SD: 1.38).

Participants’ responses to the assessment process theme showed high support for the use/scoring and design approaches. The use/scoring action (“*when planning a unit, consider when assessments and feedback will be provided to students*”) got the highest level of support from the participants (M: 5.72, SD: 0.46). All participants chose “likely” (27.8%) or “highly likely” (72.2%) options on the scale for this item.

It was followed by the design action (“*design a summative unit assessment based on what was taught to students*”) as all the participants chose “somewhat likely” (5.6%),

“likely” (33.3%) and “highly likely” (61.1%) on the scale for this item (M: 5.56, SD: 0.62). The communication action (“*discuss assignments and co-construct grading criteria for the unit with your students.*”) got the lowest score in this theme with a mean score of 3.22 (SD: 1.52).

Only the equitable fairness approach received high support from the participants in the assessment fairness theme. The equitable fairness approach (“*ensure that assessments throughout the unit are accommodated for formally identified learners*”) received the highest level of support with a mean score of 5.56 (SD: 0.62). All participants chose the “somewhat likely” (5.6%), “likely” (33.3%), and “highly likely” (61.1%) options on the scale. The standard fairness action (“*plan class lessons and assessments that are the same for all students and encompass the curriculum expectations*”) and the differentiated fairness action (“*give all students a diagnostic assessment at the beginning of the unit and have students use their results to select personalized appropriate learning and assessment activities*”) got a low score from the participants indicating that they did not favor these actions in this particular scenario. The standard fairness action had a mean score of 3.61 (SD: 1.94), and the differentiated fairness action had a mean score of 3.33 (SD: 1.37).

The responses to the assessment theory theme indicated high support for all approaches. The contextual theory action (“*in designing your assessments, prioritize the alignment between the assessment and the intended learning objectives*”) received the highest level of support from the participants (M: 5.72, SD: 0.46). All of the participants chose “likely” (27.8%) and “highly likely” (72.2%) on the scale for this item. It was followed by the consistent theory approach (“*work with other teachers to develop a reliable summative assessment for the unit*”) as all participants chose “somewhat likely” (16.7%), “likely” (22.2%) and “highly likely” (61.1%) on the scale for this item (M: 5.44, SD: 0.78). The balanced theory received the lowest level of support among the approaches. The balanced theory action (“*adapt assessments that have worked well with other students like yours to take into consideration the context and activities of your enacted lesson*”) received a mean score of 4.89 (SD : 1.32).

Table 4.5 presents the descriptive statistics that summarize participants' support for different classroom assessment approaches across all four scenarios.

Table 4.5. Descriptive statistics across 4 scenarios for all participants

Theme	Related approach	M	SD
Assessment Purpose	AoL	4.04	0.86
	AfL	4.50	0.81
	AaL	4.03	1.19
Assessment Process	Design	5.22	0.77
	Use/scoring	4.03	0.80
	Communication	4.67	0.83
Assessment Fairness	Standard	3.46	0.85
	Equitable	3.86	0.85
	Differentiated	3.82	1.12
Assessment Theory	Consistent	4.57	0.84
	Contextual	4.54	0.59
	Balanced	4.13	0.81

Participants showed high support for all three approaches under the assessment purpose theme. AfL received the highest level of support (M: 4.50, SD: 0.81). It was followed by AoL and AaL with mean scores of 4.04 and 4.03 (SD: 0.86, 1.19).

Consistent with the assessment purpose theme, all approaches within the assessment process theme received high levels of support from the participants. The design approach received the highest level of support with a mean score of 5.22 (SD: 0.77). It was followed by the communication approach with a mean score of 4.67 (SD: 0.83) and the use/scoring approach with a mean score of 4.03 (SD: 0.80).

Differing from the other two themes stated above, none of the approaches within the assessment fairness theme received high levels of support from the participants. The highest mean score recorded was the equitable fairness approach with a mean score of 3.86. It was followed by the differentiated fairness approach (M: 3.82, SD: 1.11) and the standard fairness approach (M: 3.56, SD: 0.85).

Lastly, all three approaches within the assessment theory theme received high levels of support from the participants. The consistent theory approach got the highest level of support (M: 4.57, SD: 0.84). It was closely followed by the contextual theory

approach with a mean score of 4.54 (SD: 0.59). The balanced theory approach got the lowest level of support among these three approaches with a mean score of 4.12 (SD: 0.81).

The participants' support for a specific classroom assessment approach in relation to the four themes of assessment (assessment purpose, assessment process, assessment fairness and assessment theory) specified within the ACAI framework is presented in Table 4.6 below. As stated before, if the mean score was 4 or above, it was interpreted as the participants supported that approach. However, if it was below 4, the approach was not supported. Their support for an approach is indicated with an X on the table.

Table 4.6. Individual Assessment Profiles of the Participants'

Participants	Assessment purpose			Assessment process			Assessment fairness			Assessment Theory		
	AoL	AfL	AaL	Design	Use	Commun.	Standard	Equitable	Differentiated	Consistent	Contextual	Balanced
P1	X	X				X	X			X		X
P2	X	X	X	X	X	X		X	X	X	X	X
P3		X		X							X	X
P4	X	X	X	X	X	X	X	X	X	X	X	X
P5	X		X	X		X				X	X	X
P6				X	X	X	X	X	X	X	X	
P7	X	X	X	X	X	X	X	X	X	X	X	
P8	X	X	X	X	X	X		X	X	X	X	X
P9		X		X	X	X		X		X	X	
P10		X	X	X	X	X		X	X	X	X	X
P11	X	X		X	X	X				X	X	
P12	X	X		X		X		X		X	X	X
P13	X	X	X	X	X	X		X	X		X	
P14	X	X		X								X
P15				X		X				X	X	X
P16	X							X		X		
P17			X	X	X	X			X	X	X	X
P18	X	X	X	X	X	X	X	X		X	X	X

According to Table 4.6, P1 supported the AoL, AfL, communication, standard fairness, consistent theory and balanced theory approaches. P2 showed support for all priority approaches across all themes except for the standard fairness approach. P3 showed support for AfL, design, contextual theory and balanced theory approaches. P4 showed support for all priority approaches across all themes. P5 showed support for AoL, AaL, design, communication and consistent theory, balanced theory and contextual theory approaches. P6 did not show significant support for any priority approaches within the assessment purpose theme, but she showed support for all of the priority approaches within the remaining themes except for the balanced theory approach. P7 showed support for all of the priority approaches across all themes except for the balanced theory approach. P8 supported all approaches except for the standard fairness approach. P9 supported AfL, design, use/scoring, communication, equitable fairness, and contextual theory and consistent theory approaches. P10 supported all of the priority approaches except for AoL and standard fairness approaches. P11 supported AoL, AfL, design, use/scoring, communication, and consistent theory and contextual theory approaches. P12 showed support for AoL, AfL, design, communication, equitable fairness and all three of the theory related approaches. P13 showed support for all approaches except for the standard fairness, consistent theory and balanced theory approaches. P14 supported AoL, AfL, use/scoring and balanced theory approaches. P15 supported design, communication, and all three theory related approaches. P16 supported AoL, equitable fairness and consistent theory approaches. P17 supported the AaL approach, the balanced theory approach and all three approaches within the assessment process and assessment theory themes. P18 supported all approaches except for the differentiated fairness approach.

4.2. Teaching Experience and Support for Assessment Approaches

This section presents the findings related to the second research question, which was *“How does teaching experience shape in-service EFL teachers' support for a specific classroom assessment approach in relation to the four themes of assessment (assessment purpose, assessment process, assessment fairness, and assessment theory) specified within the ACAI framework?”*. Descriptive statistics were

calculated separately for two experience-based groups named moderately experienced (5-10 years) and highly experienced (11-15 years) to answer this question. These groups will be referred to as MET (moderately experienced teachers) and HET (highly experienced teachers) throughout the chapter. Each group's preferences were analyzed independently and then compared with each other to identify meaningful differences or tendencies.

Table 4.7. Comparison of MET and HET

Theme	Approach	Moderately experienced group		Highly experienced group	
		M	SD	M	SD
Assessment Purpose	AoL	4,11	0,88	3,97	0,88
	AfL	4,61	0,70	4,39	0,94
	AaL	4,22	1,16	3,83	1,27
Assessment Process	Design	5,11	0,74	5,33	0,82
	Use/scoring	4,19	0,48	3,86	1,04
	Communication	4,75	0,70	4,58	0,98
Assessment Fairness	Standard	3,86	0,88	3,06	0,62
	Equitable	3,97	0,80	3,75	0,92
	Differentiated	4,14	1,24	3,50	0,94
Assessment Theory	Consistent	4,86	0,74	4,28	0,88
	Contextual	4,61	0,67	4,47	0,52
	Balanced	4,22	0,49	4,03	1,06

As it can be seen in Table 4.7, The MET group showed high levels of support for all three approaches within the assessment purpose theme. The AfL approach received the highest level of support from the participants (M: 4.61, SD: 0.70). It was followed by AaL (M: 4.22, SD: 1.16) and AoL (M: 4.11, SD: 0.88). Similarly, all three approaches within the assessment process got high levels of support. The design approach was recorded to have the highest level of support among the three (M: 5.11, SD: 0.74). The communication approach got the second-highest score within the theme with a mean score of 4.75 (SD: 0.70). The least supported approach was the use/scoring approach with a mean score of 4.19 (SD: 0.48). Differing from the first two themes, one of the approaches within the assessment fairness theme got high support from the participants. The differentiated fairness approach got the highest level of support with a mean score of 4.14 (SD: 1.24). Despite its high mean score, the standard deviation demonstrates that the participants had different views

about this approach. The standard fairness (M: 3.86, SD: 0.88) and the equitable fairness approaches (M: 3.97, SD: 0.80) could not reach the 4.0 threshold, so they were not supported by the participants. Finally, all the approaches were supported by the participants within the assessment theory theme. The consistent theory approach received strong support from the participants (M: 4.86, SD: 0.74). The contextual theory approach ranked second in terms of participant support (M: 4.61, SD: 0.67). The balanced theory approach was the least supported

The HET group supported one of the approaches within the assessment purpose theme, which was the AfL approach. The AfL approach got the highest level of support with a mean score of 4.39 (SD: 0.94). The AoL approach (M: 3.97, SD: 0.87) and the AaL approach (M: 3.83, SD: 1.27) remained below the established threshold of 4.0. In the assessment process theme, two of the approaches received strong support from the participants. The design approach got the highest level of support with a mean score of 5.33 (SD: 0.81), and it was followed by the communication approach with a mean score of 4.5 (SD: 0.98). The least supported approach was the use/scoring approach (M: 3.86, SD: 1.04). None of the approaches within the assessment fairness theme were reported to get strong support from the participants. All of them fell below the 4.0 threshold. The equitable fairness approach got a mean score of 3.75 (SD: 0.92), the differentiated fairness approach got a mean score of 3.50 (SD: 0.94), and the standard fairness approach got a mean score of 3.06 (SD: 0.62). Finally, in contrast to the assessment fairness approach, all the approaches within the assessment theory theme were supported by the participants. The contextual theory approach got the highest level of support with a mean score of 4.47 (SD: 0.52). The consistent theory approach came next with a mean score of 4.28 (SD: 0.88). The balanced theory approach was the least supported approach by the participants with a mean score of 4.03 (SD: 1.06).

The comparison of moderately and highly experienced teachers revealed some similarities and differences. Both groups supported the AfL approach within the assessment purpose theme. In addition to the AfL approach, MET also supported the AoL (M: 4.11) and the AaL (M: 4.22) approaches. In terms of assessment process, they both showed strong support for the design and communication approaches.

MET supported the use/scoring approach along with the design and communication approaches. As for the assessment fairness theme, standard and equitable fairness approaches did not reach the 4.0 threshold in either group. However, the differentiated fairness approach was supported by the MET group (M: 4.14, SD: 1.24). Lastly, all three approaches were supported by both groups within the assessment theory theme. Yet, the overall support for the assessment theory theme is higher in the MET group (MET M: 4.56, HET M: 4.26).

Overall, the results revealed that the participants showed varying support for different classroom approaches in relation to the four themes of assessment specified in the ACAI framework. The participants showed the highest level of support for the assessment process theme, specifically for the design and communication approaches. Also, they showed a high level of support for the assessment theory theme with consistent and contextual theory approaches receiving more support than the balanced theory approach. Within the assessment purpose theme, the participants favored the AfL approach over the AoL and AaL approaches even though they were all supported to some extent. In contrast, they did not support any of the approaches within the assessment fairness theme. In addition, the comparison of findings based on teaching experience revealed some similarities and differences between the MET and HET groups. For example, MET showed broader support for all assessment themes and related priority approaches, and they favored the differentiated fairness approach. On the other hand, HET were more selective as they supported the AfL approach in the assessment purpose theme, all three approaches in the assessment theory theme, design and communication approaches in the assessment process theme. All of these findings demonstrate that both groups have a formative assessment oriented and theory driven approach. However, the MET seem to be more open to a range of assessment approaches, especially in purpose and fairness themes.

4.3. In-service EFL Teachers' Perceptions and Experiences Regarding Classroom Assessment

This section answers the third research question, which was *“How do in-service EFL teachers perceive classroom assessment and reflect on their practices regarding*

it?”. It presents qualitative data obtained from semi-structured interviews conducted with in-service EFL teachers. Interview recordings were transcribed, and each interview was analyzed individually using inductive coding. The recurring themes were identified and named. Then, subthemes were determined to focus on specific aspects under the main theme. The themes and subthemes are supported directly with excerpts from interviews. Themes and subthemes are broadly aligned with the ACAI framework including assessment purpose and use, assessment practices, fairness and challenges and professional development needs. Detailed information regarding the themes and subthemes can be found in Table 4.8. The description for each subtheme is included in the codebook in the appendices (see Appendix F).

Table 4.8. Themes and Subthemes Derived from In-service EFL Teachers’ Interview Responses

Theme	Subtheme	Frequency (by person)
Assessment Purpose and Use	Using assessment to understand needs and track progress	17
	Using assessment as a learning tool	7
	Assessment guiding instructional planning	14
Assessment Practices	Summative assessment dominating teaching practice	6
	Preference for regular written exams	3
	Preference for alternative assessment	15
	Teacher autonomy in formative assessment decisions	15
	Hesitation toward alternative assessment	6
	Confidence in using alternative assessments	13
	Assessment outcomes shaping instruction	8
Assessment fairness	Efforts to ensure fairness in assessment	14
	Transparency in assessment criteria	5
Assessment Challenges	Time constraints	13
	Institutional/policy constraints	14
	Negative impact on teacher well-being	7
Need for professional development	Need for hands-on and collaborative training	15

4.3.1 Perceptions and Experiences Regarding Assessment Purpose and Use

The first theme that was identified was assessment purpose and use. Participants’ responses revealed three subthemes here, which were using assessment to understand

needs and track progress, using assessment as a learning tool, and assessment guiding instructional planning.

All of the participants emphasized that they used assessment to identify student needs and track progress, which led to the identification of the first subtheme, using assessment to understand needs and track progress. They indicated that assessment should serve as a tool to determine what students have learnt. One participant stated:

Excerpt 1: How can we make sure whether and how much they have learnt and how can we take them to the next level individually? This is what it (assessment) means to me. (P15, HET, 15 years of teaching experience)

Some teachers emphasized that assessment helps them determine what needs to be revisited or worked further on. Some participants supported this view of assessment purpose and use in the following excerpt:

Excerpt 2: So, first of all, we monitor the student progress, we identify the learning gaps, and also, we guide instruction. It also helps both the students and teacher to find out their understanding of what has been taught and focus more on what needs to be worked on. (P16, HET, 14 years of teaching experience)

Excerpt 3: The role of assessment is to help to understand student needs and deficiencies, to be able to help with these deficiencies, to be able to learn what topics we need to help with. (P1, MET, 9 years of teaching experience)

Another participant stated that summative assessment helps teachers identify the topics that need to be revisited in the following excerpt:

Excerpt 4: As the year goes by, whenever you have a summative assessment, you can actually see which topics are learned, and we are ready to move on. You can actually see which topics need more working on and maybe reteaching. (P5, MET, 5 years of teaching experience)

Some of the participants indicated that they viewed assessment not only as a way of keeping track of student progress but also as part of and supportive of the learning process. This led to the second child code in the category, which was using assessment as a learning tool. One participant explained it as follows:

Excerpt 5: *We are using assessment as a tool for learning more than, not just for grading. (P15, HET, 14 years of teaching experience)*

Some teachers also stated that assessment helps students keep track of their own learning progress. One participant explained it in the following excerpt:

Excerpt 6: *I would like to sit down with each individual student after the first assessment that we give to them and create a list of goals and then have like a rubric that we create together and a checklist... so they can actually see that they're making progress. (P14, HET, 15 years of teaching experience)*

Similarly, another participant emphasized the importance of enabling students to track their own progress:

Excerpt 7: *Right now, I heavily concentrate on the process of students' improvement, not their results, not their scores. I provide timely feedback. I help them track their own progress. (P10, HET, 13 years of teaching experience)*

Assessment practices that monitor and support individual student development were also frequently emphasized. One participant stated the following:

Excerpt 8: *Portfolios are great if they can cover the whole year and make the student an independent learner by monitoring improvement in each student's work. (P13, HET, 15 years of teaching experience)*

Similarly, one participant emphasized the contribution of peer assessment to learning and raising students' awareness:

Excerpt 9: *Peer assessment is very useful because when you have the right criteria and rubric, students actually can assess each other's work very well. And when they do that, I think it's one of the most effective ways to make a student understand where they're making a mistake. (P5, MET, 5 years of teaching experience)*

In addition to the first two uses/purposes of assessment, 14 participants highlighted that assessment guided their instructional planning. They said that they adapted their teaching or planned their activities and lessons according to assessment requirements. Some participants explained it as follows:

Excerpt 10: *To be honest, it definitely affects our lesson planning. In fact, we are trying to plan our lessons according to assessment. (P17, HET, 13 years of teaching experience)*

Excerpt 11: *It definitely affects lesson planning because it is necessary to make the right decisions. and decide whether it is appropriate for the upcoming lessons after applying it by updating, removing or adding new things. (P11, HET, 14 years of teaching experience)*

In a similar way, another participant drew attention to lesson planning based on assessment:

Excerpt 12: *First of all, we focus on assessment, what we want to assess, and then we start planning the lesson according to what we want to assess. (P16, HET, 14 years of teaching experience)*

Some participants stated that assessment not only affects their current lesson planning, but it also affects their lesson planning in the following year. One participant expressed it as follows:

Excerpt 13: *In fact, I think that every assessment we do should be reflected in the lesson plans throughout the year. In this sense, of course, it affects our plan and the plans for the following year, what went right, what went wrong. For these children, for children with special needs, for slow learners, for faster learners, what went right, what went wrong. What can we do after all this? Are there things we can change or adapt in the process within the same year? More importantly, what lessons did we learn from this and how can we change it in the following year? I think this is the goal. (P15, HET, 15 years of teaching experience)*

Furthermore, some participants noted that their lessons were shaped according to the data they obtained from assessment tools. One participant said:

Excerpt 14: *I plan my lessons according to the data that I get, and I prefer planning my next lesson based on this. (P9, MET, 6 years of teaching experience)*

Finally, some teachers emphasized the importance of identifying the strengths and weaknesses of learners and reflecting on these in the lesson planning process. Some participants said the following:

Excerpt 15: *I can identify students' strengths and weaknesses. I can consider re-teaching the topic, or let's say to just move forward. (P4, MET, 10 years of teaching experience)*

Excerpt 16: *With assessment, there are many times when I plan my own lesson by looking at what the students did under those rubrics. (P3, MET, 6 years of teaching experience)*

In conclusion, the participants expressed that they viewed assessment as a multifunctional tool that serves many instructional purposes. The findings showed that participants use assessment in three main ways: tracking student progress and identifying needs, supporting the process of learning and guiding instructional planning. The participants often stated that they used the data they got from the assessment to identify what needs to be worked on further and what needs to be adapted or removed accordingly. Also, they said that assessment affects not only the current teaching practices but also long-term planning and decision-making processes.

4.3.2. Perceptions and Experiences Regarding Assessment Practices

In addition to assessment purposes and use, the participants also shared their perceptions and experiences related to assessment practices. This led to the identification of the second theme named assessment practices. Under this theme, the participants touched upon various issues such as exam preparation, alternative assessment tools and teacher autonomy in designing and implementing summative and formative assessments. The first subtheme within this category is summative assessment dominating current practice.

6 participants indicated that they taught or planned according to the requirements of the summative assessment. They stated that they relied on summative assessment-based practices to prepare their students for these exams. One participant said the following:

Excerpt 17: *We prepare them (the students) for summative assessment with additional materials in the classroom, and we have to do it like this.*

Otherwise, the students will be shocked when they see the exam. (P17, HET, 13 years of teaching experience)

Similarly, other participants also emphasized how much it affects their teaching in this way:

Excerpt 18: It's evolved in that sense... and now it's a lot more preparation, kind of teaching towards the test, even if I haven't seen the test. (P7, MET, 6 years of teaching experience)

Excerpt 19: I am constantly preparing the students for a quiz or an exam, and it is as if I am not doing what the curriculum wants me to do, but what that exam requires. At some point, our main focus becomes the exam. And this has a serious backwash effect on me as a teacher. Even when I give an example in class, I find myself saying that it may appear in the exam. And this is not a pleasant situation for me. (P3, MET, 6 years of teaching experience)

Some participants stated that high stakes or national examinations also affect teaching practices. Some participants stated the following:

Excerpt 20: In the end, there is this reality in our country and in the whole world that you need to pass some certain exams in order to be considered competent. And there is no differentiation in these exams. (P15, HET, 15 years of teaching experience)

One teacher said that summative assessment leads to changes and adaptations not only in the way that she teaches but also in what she teaches:

Excerpt 21: For example, focusing more on some specific topics for the achievement or proficiency tests may lead to changes in course content as well. (P13, HET, 15 years of teaching experience)

Moreover, one participant stated that many students studied not for learning but for the summative assessment:

Excerpt 22: Many students are not really studying to learn but more for assessment. (P6, MET, 7 years of teaching experience)

One teacher commented on how institutional factors contributed to summative assessment dominating the practice:

Excerpt 23: *So actually, the system pushes us a little bit more towards traditional methods, and it's a little bit easier for us to get that data through traditional methods. (P9, MET, 6 years of teaching experience)*

Finally, two participants drew attention to the pressures put on both teachers and students due to the frequency and weight of summative assessments. They said:

Excerpt 24: *Sometimes I feel like our students get exposed to a lot of summative assessments in school. It's difficult, of course, for both students and teachers as well because we have to evaluate and give feedback all the time. (P4, MET, 10 years of experience)*

Excerpt 25: *Since some institutions in Turkey still emphasize summative assessment, the teacher's style inevitably changes accordingly. (P3, MET, 6 years of teaching experience)*

The next subtheme was preference for regular exams. This subtheme was used only by 3 moderately experienced teachers. However, the researcher chose to include it because it relates to the diverse views among MET and HET. One participant stated that written exams were best for teachers who were at the early stages of their careers. She said the following:

Excerpt 26: *I think many teachers feel safer here. As a teacher at the beginning of my profession, I also feel safer because you have a rubric in front of you, a checklist. You look according to it, where there are problems and where there are not. You have a very clear picture in front of you. (P9, MET, 6 years of experience)*

Another participant also reported that she preferred traditional exams. She commented:

Excerpt 27: *Well, I prefer the regular exams because it gives you more detailed results, of course. (P7, MET, 6 years of teaching experience)*

Finally, one participant stated that she trusted traditional exams more than the alternative methods in the following way:

Excerpt 28: *To be honest, I may be a bit old school here, but I think traditional assessment methods are more effective. (P6, MET, 7 years of teaching experience)*

In contrast to preference for regular exams, the majority of the participants preferred alternative assessment methods. These teachers reported that they think alternative assessment methods, such as self and peer assessments, portfolios, projects, and observations contribute greatly to learning and improvement. One participant explained her strong preference for alternative methods in the following way:

Excerpt 29: Well, I am all for formative assessment. I think it has more impact on students metacognitively, you know. Assessment, I mean using formative assessment in the classroom, is my favorite. It's what we all love. (P17, HET, 13 years of teaching experience)

Another participant said that students could perform better in flexible and interactive environments. That's mainly why she preferred alternative assessment methods. She explained this in the following way:

Excerpt 30: When I see a student's performance in the classroom, I understand that s/he can perform better in flexible environments. In face-to-face environments...However, s/he cannot in the exam where they sit with a paper for a certain period of time and cannot establish a connection with his/her teacher and friends. I think this is a big issue. (P15, HET, 15 years of teaching experience)

One participant stated that alternative assessment was in the nature of teaching, and she used it naturally. She expressed her opinion in this way:

Excerpt 31: Alternative assessments are always there because peer-assessment, self-assessment, group work, exit tickets... These are things that we always do. This should be a part of the lesson. (P11, HET, 14 years of teaching experience)

Furthermore, some participants shared their opinions about the most effective alternative assessment tools. Two participants stated that they preferred portfolios and projects among alternative assessment methods:

Excerpt 32: But in general, throughout the school year, I would rather see things like projects and portfolios. (P14, HET, 14 years of teaching experience)

Excerpt 33: *Mostly when students have to work on a project together, I think that's the most effective one because they can remember many things from them. (P5, MET, 5 years of teaching experience)*

Two participants said that they preferred self and peer-assessment among alternative methods:

Excerpt 34: *But for me, I think the most important one is self-assessment. And I really like the exit tickets because it helps me to collect the data very well and to see which students understood wrong or correctly. (P4, MET, 10 years of teaching experience)*

Excerpt 35: *Alternative assessment, like in-class portfolios, like peer assessment, self-assessment, etc. These are quite valuable. I was going to add these. (P10, HET, 13 years of teaching experience)*

One participant focused on the importance of classroom discussions and explained her preference for the think-pair-share technique:

Excerpt 36: *Think-pair-share is one of my favorites. I especially like it when we have discussions. That's why I like think-pair-share because it also turns into discussion after a while. That's also quite nice. I think I like it the most. (P11, HET, 14 years of teaching experience)*

Finally, some participants stated that they prefer a blended approach because they like to use both alternative and traditional assessment methods and tools. One participant explained it in this way:

Excerpt 37: *But if I need to give like a general answer for that, I can say I prefer a mix of traditional and alternative assessment. (P10, HET, 13 years of teaching experience)*

The fourth subtheme was teacher autonomy in formative assessment decisions. This subtheme emerged because 15 participants talked about being able to redesign or adapt the formative assessment tools that were included in their lesson plans. One teacher explained this as follows:

Excerpt 38: *We are all following the same lesson plan, but we can be more flexible here within levels. We do not necessarily have to implement the same*

formative assessment tool in our classes. After designing a tool that serves our purpose and sharing it with our direct-level colleagues and getting approval, I think we can use it at this point. We are already doing this. (P9, MET, 6 years of teaching experience)

One participant stated that she even made on-the-spot changes and adaptations whenever needed. She said the following:

Excerpt 39: In line with the need I realize in the classroom, I design it right away and always apply it in formative assessment. Sometimes I even cancel my plan and decide that the priority of my class is to go through this formative assessment structure and apply it in the class. (P15, HET, 15 years of teaching experience)

Some participants emphasized the flexibility that the institution provides them with when it comes to formative assessment decisions. They stated the following:

Excerpt 40: Yes, we do have the freedom to do formative assessments. And whichever type formative assessment we want, we can implement that in our lessons. (P5, MET, 5 years of teaching experience)

Excerpt 41: Of course, we can if we just, you know, think that there is a different assessment tool that we can use and we can suggest it and change it. So yeah, I mean, it's not that strict. (P7, MET, 6 years of teaching experience)

One participant said that making changes or adaptations in the formative assessment tools was encouraged by the institution:

Excerpt 42: There's a lot more freedom with that (formative assessment), thankfully. And it's encouraged. (P14, HET, 15 years of teaching experience)

Finally, some participants said that they viewed differentiation as an essential part of formative assessment, so they differentiated the formative assessment tools as needed.

Excerpt 43: We definitely change and differentiate formative assessments. (P11, HET, 14 years of teaching experience)

Excerpt 44: Of course, we do have flexibility and autonomy in formative assessment. We also differentiate tasks ourselves when we need to. (P7, MET, 6 years of teaching experience)

The fifth subtheme was hesitation toward alternative assessment. Similar to the other subthemes that have been covered in this section so far, this subtheme was used by both moderately and highly experienced teachers. However, unlike the other subthemes, their reasons were different from each other. 2 MET expressed that they did not feel confident and knowledgeable enough to use alternative assessment tools. They stated that:

Excerpt 45: I especially think that I need to study and work on myself a little bit more. So I think I need to work on it a little bit. (P4, MET, 10 years of teaching experience)

Excerpt 46: Likewise, when it comes to the preparation and evaluation process of these (alternative assessment tools), I personally think that I lack knowledge and feel inadequate. This also causes me to somewhat avoid using them. (P9, MET, 6 years of teaching experience)

Unlike these two MET, one MET said that she hesitated to use these alternative methods because of artificial intelligence. She explained it in the following excerpt:

Excerpt 47: Because now you know that students can use AI everywhere, so that is not a good way to assess and to know exactly how well they are doing because they can go home, and they can just use AI or use anything from the internet to create a project or a portfolio. (P7, MET, 6 years of teaching experience)

The HET had a different reason as to why they hesitated to use alternative assessments. They did not make a general comment about alternative assessments, but they focused on two alternative assessment tools, which were portfolios and projects. They explained it in the following excerpts:

Excerpt 48: I do not feel very competent in things like portfolios and projects because I haven't done much. I've done one or two, but I do not think I've done it justice. (P11, HET, 14 years of teaching experience)

Excerpt 49: I have less confidence using something like a portfolio or a project because I do not have hands-on classroom experience with it yet. (P14, HET, 15 years of teaching experience)

The sixth subtheme was confidence in using alternative assessments. In contrast to the participants who expressed hesitation toward using alternative assessments, the

majority of the participants noted that they used self-peer assessment, exit tickets, observation and felt competent in doing so. One participant explained that in the following excerpt:

Excerpt 50: Well, I do. I feel confident because I also believe that we should give students some freedom to use different kinds of assessments like self-assessments, to know how well they can also assess themselves or to make them create some projects. (P7, MET, 6 years of experience)

Another participant said the following to emphasize how confident she feels in terms of observation:

Excerpt 51: Well, I think that I am competent in observations because I am used to observing the students in the classroom somehow. As much as I understand there, I can monitor them immediately. (P17, HET, 13 years of teaching experience)

Furthermore, some participants linked their confidence to institutional culture. They stated the following:

Excerpt 52: But since we use other methods (alternative assessment tools except for portfolios and projects) very often, especially in my current organization, I feel competent in these matters. (P11, HET, 14 years of teaching experience)

Excerpt 53: Alternative ones. I feel competent because I am aware that in my current school, these are actually valued more than in other schools. (P15, HET, 15 years of teaching experience)

The final subtheme under this theme was assessment outcomes shaping instruction. In addition to making autonomous formative assessment decisions, 8 participants also emphasized that they proactively used assessment results to adapt teaching.

One participant described how exit tickets are used to adapt instruction based on what has been learnt and what needs to be worked further on. She explained it as follows:

Excerpt 54: Exit ticket is also an important formative assessment method for us in our lesson plans. We receive feedback on what and how much the students have learnt at that moment. We receive it so that we can make an

arrangement accordingly. We actually make improvements. (P17, HET, 13 years of teaching experience)

The same participant also added:

Excerpt 55: We understand the problem there, and we revise it with questions similar to the problematic items. (P17, HET, 13 years of teaching experience)

In addition, some participants stated that they started preparing or shaped their lesson plans according to the outcomes of the assessment. They stated the following:

Excerpt 56: First of all, we focus on assessment, what we want to assess, and then we start planning the lesson according to what we want to assess. (P16, HET, 14 years of teaching experience)

Excerpt 57: I think every assessment needs to be analyzed and synthesized by the teacher when the assessment is over; and I think that every assessment we do should be reflected in the lesson plans throughout the year with this in mind. (P15, HET, 15 years of teaching experience)

Finally, some participants mentioned that they decided whether the topic needed to be revisited based on assessment results as well. They explained this in the following excerpts:

Excerpt 58: It's to find out what they know and what they've learned and what needs to be retaught, reviewed, or if they've mastered it and you can just move on. (P14, HET, 15 years of teaching experience)

Excerpt 59: As the year goes by, whenever you have a summative assessment, you can actually see which topics are learned and we are ready to move on. You can actually see which topics need more working on and maybe reteaching. (P5, MET, 5 years of teaching experience)

To summarize, the participants shared various experiences and perceptions regarding assessment practices under this theme. They stated that these practices are shaped both by individual approaches and institutional factors. They indicated that summative assessments dominated their instructional practices because they needed to get students prepared for them. The preference for traditional pen and paper exams has been mentioned only by 3 MET, which showed that the use of alternative assessments is not only related to the objectives of the lesson, but it is also related to

professional confidence. In line with this, some teachers expressed lack of confidence in using alternative assessments due to a lack of knowledge or confidence. Conversely, most of the participants expressed confidence in using various alternative assessment tools. Moreover, the majority of participants said that they had autonomy in differentiating, adapting or redesigning formative assessments whenever needed. Finally, it was reported that both summative and formative assessment practices affect lesson planning and instruction in terms of lesson content, progress and revision of specific topics or items. All of these findings show how teachers assess their students is shaped by many factors such as context and experience.

4.3.3. Perceptions and Experiences Regarding Assessment Fairness

In addition to assessment purpose, use and practices, the participants also shared their perceptions and experiences related to assessment fairness. This led to the identification of the third theme, assessment fairness. Under this theme, the participants talked about adapting, interpreting or designing assessments in various ways and sharing assessment criteria openly to ensure fairness. Participants mentioned various methods they used to promote fairness among students while assessing their performance. Two subthemes were identified under this theme. The first one was efforts to ensure fairness in assessment. Under this subtheme, 14 participants talked about various efforts to ensure fairness among learners such as considering individual needs, using differentiated tasks and tools, talking with colleagues and maintaining consistency across assessments. One participant stated that in order to be fair, each student should be assessed based on his/her own developmental process and in-class participation. She explained that in the following excerpt:

Excerpt 60: In order to ensure fairness in the classroom, that is, not in the classroom, but in assessment, I try to use teacher assessment grades a little bit in my own way. Because when you look at the child's performance in the classroom and when you look at the child's performance in the exam, you really say that we have many such children. I mean, I think that the child's participation in the class, his/her ability to speak English, his/her self-development, each child's development is his/her own, so the child does not

necessarily need to get a score above 85 in order to be considered good. (P17, HET, 13 years of experience)

Similarly, another participant expressed that she supported her students that studied hard through the teacher assessment grade:

Excerpt 61: I try to give them extra points from teachers' assessment, you know, because I see that they are really trying. (P5, MET, 5 years of teaching experience)

One teacher stated that she used differentiated assessment tools to ensure fairness in assessment. She said the following:

Excerpt 62: So, for example, when it comes to formative assessment, especially we try to design this assessment in a way to differentiate according to the students' interests, learning or something like that. So, I think in this one, yeah, we are pretty fair. (P4, MET, 10 years of teaching experience)

Two participants mentioned that differentiation helps them make assessments fairer:

Excerpt 63: As I said, as long as I use differentiation as actively and effectively as possible, I think it is a little bit fairer in my eyes. (P6, MET, 7 years of teaching experience)

Excerpt 64: We differentiate. I think when you get to know the children and address their needs, we are actually being fair. (P14, HET, 14 years of teaching)

One participant emphasized the importance of open communication among colleagues in order to maintain fairness. She explained it in this way:

Excerpt 65: Personally, I'm trying my best to be fair as much as I can, to also speak with my colleagues to discuss some of the answers. So, at least we can be on the same page altogether to have like a fair assessment. (P7, MET, 6 years of teaching experience)

The same participant also added the following to indicate that they show a group effort in order to be fair to all students:

Excerpt 66: For example, when we are choosing the audio for listening exams, we really try our best to make it heard and understandable for all the

students, no matter what their level is. (P7, MET, 6 years of teaching experience)

Finally, some teachers expressed that they follow standardization and double grading procedures in summative assessments to make sure that assessments are evaluated fairly by all teachers. One participant explained these procedures in the following excerpt:

Excerpt 67: First of all, we do standardization. We give points to the same recordings from all schools, not just our Ankara school, but from all schools. According to the rubric, according to the evaluation criteria. Then, our coordinator examines the scores we give. And finally, a decision is reached. And then we are given a standardized scoring range for each recording or paper. There is also something like this in grading writing. We do second marking on the writings. The first marker grades the papers. S/he gives the papers to the second marker without sharing the scores. The second marker grades and again does not share his/her scores. Finally, they come together and compare the scores and if the difference is greater than two points, it goes to the third marker. If there is a situation that cannot be reconciled with the third marker, we get support from our coordinator, our department head. So, this is an effort to maintain a little bit more credibility, to be standard and to be fair. I support this. (P6, MET, 7 years of teaching experience)

The second subtheme was transparency in assessment criteria. Under this subtheme, 5 participants emphasized the importance of sharing clear and reliable rubrics openly with the students. One participant highlighted that the assessments should be structured and based on rubrics in the following excerpt:

Excerpt 68: Of course, you know, somehow, of course, all of these (assessment practices) should have something like this. I mean it has to be based on a rubric. Yes, it needs to be structured. (P17, HET, 13 years of teaching experience)

Some participants stated that the students should see the rubric or the criteria before exams or even have access to them during the exam. They stated the following:

Excerpt 69: They should actually be seeing these rubrics beforehand, to be honest. Honestly, they should have the rubric next to them while they're taking the test. (P14, HET, 15 years of teaching experience)

Excerpt 70: Establishing rapport, involving students in the process and et cetera. Yeah, this is how I do that (ensuring fairness). Sharing objectives with

the students and rubrics and stuff. This is what I mean by involving students in this process. (P10, HET, 13 years of experience)

Similarly, another participant added the following:

Excerpt 71: Because we upload the rubric in a very transparent way, we upload our criteria. Students actually give feedback to each other (in speaking exams). (P6, MET, 7 years of teaching experience)

Finally, one participant emphasized the importance of using valid and reliable rubrics for fair assessment. She said the following:

Excerpt 72: It (fair assessment) means using a valid rubric and deciding wherever you need to take points or add points, you know, for all students, it shouldn't be different for any of the students. (P5, MET, 5 years of teaching experience)

In conclusion, these findings revealed that participants used different strategies to ensure fairness in assessment. Common use of rubrics, openly sharing the assessment criteria and involving students in the process actively were among these strategies. The participants think that these help students understand the requirements clearly and make the assessment processes fairer.

4.3.4. Perceptions and Experiences Regarding Assessment Challenges

In addition to discussing how the participants use and design assessments and what kind of strategies they use to make assessment fair, they also talked about the challenges that they faced while implementing classroom assessments. This led to the identification of the third theme, assessment challenges. Under this theme, the participants mentioned challenges that stemmed from contextual or systemic factors. Three subthemes were identified under this theme. The first one was time constraints. The majority of the participants talked about time limitations being one of the biggest challenges that they faced while designing and implementing and acting on assessments. One participant explained it in this way:

Excerpt 73: The first one, I think all teachers suffer from this one, time management. So, preparing the assessment, implementing it, providing timely

feedback for students and staff. This is why time is also a great concern for students at the same time. Sometimes we do not even provide them with enough time to get feedback, spend some time on it, and then act accordingly. (P10, HET, 13 years of teaching experience)

Excerpt 74: And I think sometimes we have problems that time does not allow us to use all of them (different assessment tools) as much as we want. (P16, HET, 14 years of teaching experience)

Another participant explained that grading and analyzing assessment while continuing with her daily teaching activities affected her negatively because she did not have enough time to do all of them:

Excerpt 75: There is pacing, and we continue to teach. Let's say I graded the exam, analyzed the results and entered them into the system. What am I going to do next? There is neither time nor energy to do anything about it. (P9, MET, 6 years of teaching experience)

Another participant focused on time management problems due to the implementation of assessment. She explained that as follows:

Excerpt 76: Typically, what it (implementing assessments) does is it shortens the time that I have (in the classroom) and makes pacing more difficult. (P14, HET, 15 years of teaching experience)

The second subtheme under this category was institutional/policy constraints. This subtheme was used by 14 participants to refer to the top-down policy and the centralized decision-making process of the institution when it came to summative assessments. Previously, the participants reported having autonomy in designing and implementing formative assessments. However, unlike formative assessment, they said that they have little to no autonomy in designing and implementing summative assessment. They explained this situation in the following excerpts:

Excerpt 77: We also have a testing unit for summative. Our testing unit prepares them. In terms of questions, I mean, we do not have autonomy in terms of detailed questions, but of course we always work in consensus. (P11, HET, 14 years of teaching experience)

Excerpt 78: I'm talking about summative assessment. We can't take initiative because we work in a corporate place. There is a unit for this. The Testing unit. They prepare the exams. (P6, MET, 7 years of teaching experience)

Despite not having a lot of control over the design of summative assessments, one participant added that they can make the necessary changes for officially identified learners.

She said:

Excerpt 79: We can design or differentiate summative assessment for our identified learners. (P5, MET, 5 years of teaching experience)

The final subtheme within this category was the negative impact of assessment on teacher well-being. This subtheme was used only by MET. Out of 9 MET, 7 of them reported that assessment and assessment-related responsibilities had a negative effect on their personal well-being and professional workload. They talked about time pressure, exhaustion and a continuous feeling of inadequacy during intensive assessment periods.

Two participants explained this situation as follows:

Excerpt 80: Until I read and analyze all the results, I go through a very tiring process, and a lot of time passes (P9, MET, 6 years of teaching experience)

Excerpt 81: Honestly speaking, sometimes we have to work until late hours when we are very stuck and busy, and here I think the students are the most affected ones. Because when we come back the next day, our energy is lower, we are more tired, we cannot rest. At this point, I think the classes are too crowded. I think we teach too many classes. Inevitably, the number of papers we grade can increase. The amount of feedback we give can increase. So, I do not think this process has a very positive effect on me (P6, MET, 7 years of teaching experience)

Two participants pointed out that the negative impact of assessment on their well-being affected the students as well:

Excerpt 82: We have to evaluate and give feedback all the time. Students get overwhelmed, they get confused with the topics and it's difficult, of course, for both students and teachers as well. (P4, MET, 10 years of teaching experience)

Excerpt 83: You feel this burden that you are not doing enough for them. You are always feeling that you can do more, but you do not have the time or the support. (P7, MET, 6 years of teaching experience)

Overall, the participants talked about various factors that affect their ability to implement assessments. It was reported that time constraints were the main limitation in terms of designing, implementing and acting upon assessments. The participants especially emphasized that the top-down approach in summative assessments prevented them from making accommodations and modifications for all of their learners. Finally, the MET talked about the negative impact of assessment on their well-being as it caused fatigue, time pressure and a constant feeling of inadequacy.

4.3.5. Perceptions and Experiences Regarding Professional Development

All of the participants talked about their strong demand for further professional development opportunities. They mainly emphasized the need for professional development activities that would support teachers in designing and implementing assessment practices. This led to the identification of the final theme, the need for professional development. Since all of the participants highlighted their desire to have practical and collaborative professional development activities, the need for hands-on and collaborative training appeared as the only subtheme under this category.

One participant expressed her opinion on professional development activities in the following way:

Excerpt 84: I think more like hands-on examples would have been more practical and applicable for middle school. (P16, HET, 14 years of teaching experience)

Similarly, another participant said:

Excerpt 85: If it is possible, it should be hands-on. Something should be tried by the teachers and supervised by experts. I think pilot groups should be formed at every level. And the things to be tried should be tried and learned by teachers within pilot groups. (P11, HET, 14 years of teaching experience)

In addition, one participant added that the training should be applicable to real classroom environments:

Excerpt 86: Teachers should be given training on assessment that is based more on exemplification and practice. These training courses should be things that teachers can apply one-on-one, and they should be designed according to the real classroom environment. (P3, MET, 6 years of teaching experience)

Finally, one participant emphasized that these professional development activities should be sustainable and long-term:

Excerpt 87: I do not think it should be a one-day or two-day program, I think it should be spread over a longer period of time. For example, it should be a process that is spread over two months, three months, and at the same time, it should be a process with both input and output in the form of a workshop. I think that something that we learn by living should definitely be considered. (P12, HET, 14 years of teaching experience)

Overall, the findings revealed that in-service EFL teachers viewed assessment as a tool to track progress, support learning and guide instruction. Despite summative assessment dominating their teaching practices, the majority of the teachers valued alternative assessment methods. However, there were some differences between the experience groups when it came to their confidence levels in using alternative assessments. Even though most of the teachers in the HET group stated that they felt confident in using alternative assessment methods, 2 of them indicated that they did not feel confident enough to use portfolios and projects due to lack of experience. On the other hand, almost half of the MET stated that they did not feel confident and knowledgeable enough to use alternative assessment methods. Also, 3 MET claimed that traditional exams were safer and more reliable.

Both groups showed effort when it came to issues related to fairness in assessment. The participants pointed out that they were trying to ensure fairness in assessment with the help of differentiation, clear rubrics, transparency in assessment criteria and consistency in scoring among teachers. Despite all these efforts, teachers mentioned some challenges that they faced while designing, implementing and analyzing assessments such as time and institutional constraints. Moreover, the majority of the MET claimed that assessment had a negative impact on their well-being.

Finally, even though some differences were observed during MET and HET regarding their experiences and opinions about assessment, both groups expressed their strong desire for practical, sustainable and collaborative professional development activities.

When quantitative and qualitative findings of this study are taken into consideration, both data show that the participants view assessment as more than a tool to measure student achievement. Regarding assessment practices, quantitative data highlighted design and structured planning as can be seen in participants' high level of support for the design approach, which referred to designing reliable assessments that measure achievement of intended learning goals. This aligned with the participants' efforts to align instruction with assessment outcomes. However, when it came to their actual teaching practices, the dominance of summative assessments persisted even though there is an increased interest in alternative methods such as peer and self-assessment. Fairness theme appeared to produce more complicated findings because while it was the least supported approach in the quantitative data, the qualitative data revealed many efforts to ensure fairness in assessment practices. In addition to all this, qualitative data provided an insight into some challenges that teachers faced in their daily teaching activities such as time and institutional constraints and their professional development needs.

CHAPTER 5

DISCUSSION

This chapter presents the discussion of the results based on the relevant literature, the implications for future educational contexts and limitations of the study.

5.1. Discussion

The current study aimed to investigate the assessment approaches of in-service EFL teachers working at a private middle school, explore whether teaching experience influenced their assessment approaches and gain insights into their understanding of classroom assessment. The participants' responses were analyzed and reported according to the four main themes of assessment identified in the ACAI framework: assessment purpose, assessment process, assessment fairness and assessment theory. (ACAI v3.0, 2020).

In the assessment purpose theme, all teachers showed more support for AfL. This finding aligned with DeLuca et al. (2018), Coombs et al. (2018), and Öz and Atay (2017), who stated that there had been strong support for AfL in their studies as well. It shows that teachers are aware of the importance of formative assessment in enhancing their teaching practice, and they value it as a learning tool. While MET supported all three purposes within the theme, the HET group was more selective. This might be due to the fact that HET have been more exposed to top-down assessment policies over the years and are more cautious about implementing assessment practices according to the policy requirements. Also, similar to Öz and Atay's (2017) study, all participants stated that they used assessment to understand students' needs, track progress, support learning and guide instructional planning in the semi-structured interviews. They highlighted the importance of using assessment

results to improve their teaching, and HET noted that the results should be used to help students reflect on their performance and set goals for themselves. This view aligns with the AaL approach in the ACAI framework. Berry et al. (2019) and Önalın and Gürsoy (2020) also mentioned similar results about using assessment to inform instruction. Also, Jannati (2015) reported that participants used the assessment results for remedial practices. Furthermore, the participants' perceptions about the purpose of assessment showed that they had a sociocultural view of assessment. As defined by Willis et al. (2013), assessment is viewed as "a dynamic and context-dependent social practice" (p. 242). The participants stated that they used assessment to improve their long-term planning, they continuously adapted or changed their assessment tools for their context, student needs and fairness concerns and used their own teacher judgment to decide whether something is useful for their classroom's needs, and they associated assessment with their well-being. As can be seen, all these practices are indicative of a sociocultural view of assessment. Even though AfL was the dominant approach among the three, AaL and AoL were also supported by the participants. This finding aligns with many studies in the literature (Mutlu, 2025; Latif & Wasim, 2024; DeLuca et al., 2018; Cooms et al., 2018), and it is consistent with the qualitative data as the participants expressed using and valuing alternative assessment tools such as self-assessment, peer-assessment and goal-setting practices for the students, which aligns with the AaL approach. Furthermore, based on what participants said in the interviews, many summative assessments have to be administered throughout the semester, which explains their tendency to adopt the AoL approach. This situation mirrors the policy-practice mismatch in Türkiye that Yeni-Palabıyık and Dalođlu (2025) mentioned in their study. The fact that teachers are willing to use AaL methods such as peer-assessment, self-assessment, portfolios and projects shows that teachers are making the effort to involve students in their own learning journey even within a restrictive and clearly defined testing culture that is dominant in the research setting. When compared to other exam driven contexts mentioned earlier such as Iran, Bangladesh and China, the teachers in the current study demonstrate more agency and autonomy to indicate formative assessment tools because as they stated in the interviews they were encouraged to implement these practices by the institution. It indicates that private school contexts may provide slightly more flexibility for formative student-centered practices despite the

dominance of summative assessments. Also, this combination of formative and summative assessment practices and the dominance of the AfL approach may indicate a transition within the institution's assessment policy to a more formative assessment-oriented one.

In the assessment process theme, similar to Yastibaş and Takkaç's (2018) study, the design approach got the highest level of support from all of the participants. Therefore, it can be said that teachers give a lot of attention to designing valid and reliable assessment practices that would serve their instructional goals. Furthermore, since the teachers' lessons are regularly observed and they are given constant feedback regarding the assessment activities that are included in the lesson plans that they prepare, their personal assessment approach might have shifted to a more design-oriented one over the years. Another approach that also got a comparatively high score was the communication approach. Coombs et al. (2018) also reported similar results about adopting the communication approach in all career stages. The communication approach focuses on transparent and feedback-oriented dialogue between teachers, students and parents (DeLuca et al., 2018). All teachers supporting this approach means they are aware of the importance and impact of communicative clarity in assessment on self-awareness both in students and in parents. Their emphasis on communication may reflect the expectations of private school contexts, where all stakeholders demand clarity and accountability. Use/scoring approach, which is related to modifying rubrics or scoring practices to address assessment-related issues (DeLuca et al., 2018), got the lowest support among the three approaches under this theme. This might be because of the standardized summative nature of the institution's assessment policy, which contains procedures such as blind double grading, spot-checking and standardization. All these procedures might restrict the teachers' ability to modify rubrics or scoring practices. Furthermore, the scoring approach was only supported by the MET. Thus, while the MET group seems to be more flexible in adjusting their scoring practices and rubric in certain circumstances if given the opportunity, the HET group chooses to comply with the requirements of standard scoring procedures. However, the qualitative findings contradict this view. According to the qualitative data, it can clearly be seen that almost all of the teachers differentiate, design or adapt formative

assessment tools based on their students' needs. The reason why this difference between quantitative and qualitative data emerged might be because the participants might have compartmentalized their autonomy by complying with the constraints of standardized summative assessment practices and using the flexibility that they have in formative assessment practices. Some of the participants (especially the HET) might have focused on their own summative assessment practices while deciding whether to support this approach in scenario-based questions. It may also be because HET might have developed a more requirements-oriented approach over the years when it comes to their scoring practices, but MET might still be establishing their own professional identities and experimenting to find the best practices.

In the assessment fairness theme, none of the approaches were supported by the teachers. This finding does not align with some of the previous studies conducted in North American contexts. For example, DeLuca (2018) and Coombs et al. (2018) stated that the participants in their studies showed high support for especially for differentiated fairness. Yet, it aligns with many studies conducted in the Turkish context such as Mutlu (2025) and Çimen (2022). Similar to the teachers in the current study, their participants also reported their concerns regarding fairness in assessment due to the dominance of summative assessments and their inability to modify or adapt them based on their students' needs and preferences. This might show that the dominance of summative assessment is stronger in the Turkish context, and this affects teachers' understanding and practices of fairness in assessment. However, when the responses were analyzed by experience groups, it was revealed that the differentiated fairness approach was supported by the MET group. Qualitative findings explained how teachers understand and apply fairness in practice, but their perception apparently does not match their practices. This may suggest "a terminology awareness gap and limited theoretical framing" as it was reported by Berry et al. (2019). Both groups reported using many strategies such as considering the students' in-class performance and developmental progress while assigning teacher assessment grades. Furthermore, many participants said that they used differentiated formative assessment tools which they have designed or adapted based on their students' needs and interests. Some participants indicated that they used teacher assessment grades to support the students' overall grades as an act to

humanize assessment and support individual student learning. Even though the third strategy may seem like an attempt to make assessment fairer, it actually shows that teachers are also aware of some of the “unfair” assessment procedures, and they are trying to support their students through their own efforts. Several participants mentioned collaborative efforts as a fairness strategy. They described their standardization and double grading procedures to ensure reliability and consistency among teachers in terms of scoring. One participant explained these procedures in detail, and several others mentioned it as a strategy that greatly contributed to fairness in assessment. Although this view aligns with the ACAI framework’s emphasis on fairness as a collective effort, it also proves that the participants are uncertain about the definition of fairness, and they associate it with consistency and equality. This finding shows that fairness in assessment is understood in a superficial sense, and the teachers’ understanding of fairness has been greatly influenced by the institution’s attempts to unify assessment practices across campuses and increase exam reliability. Finally, being transparent with the students has also emerged as a fairness strategy. Some participants stressed the importance of sharing the assessment criteria and rubrics clearly with the students, and they connected this with the communication approach of the ACAI framework by saying that the expectations and the grading procedures should be clearly explained to the students through rubrics and the criteria. This finding shows parallelism with Jannati’s (2015) study, as the participants in her study also argued that teachers should inform the students about what they will be evaluated on. It can be said that even though there are efforts to promote fairness in assessment, these efforts remain unbalanced and limited by institutional constraints.

In the assessment theory theme, all of the approaches were strongly supported by the participants. The highest level of support was given to consistent theory and contextual theory. This shows that teachers emphasize procedural dependability and contextual applicability. This finding contradicts with Doğan and Ünal’s (2024) study, which showed competence gaps due to lack of theoretical knowledge, the participants of this study showed high levels of theoretical awareness regarding assessment processes. This might be due to the emphasis on assessment processes within the institution’s assessment policy and regular assessment training that are

provided by the institution. When the results are analyzed by groups, it is revealed that there is not much of a difference between their approaches to assessment theory. These findings were backed up by qualitative findings, too. During the interviews, teachers regularly referred to assessment as something integrated into teaching rather than an additional tool. Most of them talked about how assessment regularly shaped their lesson plans and how they made use of the feedback they got from the assessment tools to improve their short and long-term planning. They specifically said that assessment outcomes helped them decide whether to reteach, review, or progress with the next instructional goal. Moreover, many participants emphasized the importance of using clear rubrics and scoring guidelines, which showed alignment with consistent theory. However, balanced theory received relatively low support from the participants, indicating that they did not openly state that they should have a balanced approach of consistency and contextual considerations. This may be because the participants did not receive sufficient training or support in balancing fairness, validity, and reliability, which is strongly promoted by the balanced theory approach within the framework.

In addition to their responses to the questionnaire, the participants' answers to the interview questions revealed some valuable information about institutional constraints, teacher autonomy and professional beliefs. Whenever the participants talked about their teaching practices, they felt the need to refer to summative assessments. According to many participants, the amount and weight of summative assessments led to their practices being dominated by summative assessments, especially during certain periods throughout the term. This common practice of teaching for the test shows the institutional pressure that requires AoL to be dominant in practice, even though the teachers make a lot of effort to integrate more formative assessments into their teaching practices. When the research setting is considered, this finding is quite expected because the school has a centralized testing unit that limits teachers' autonomy in designing, modifying and adapting summative assessments. This clearly indicates that there is an urgent need for institutional reform for a more holistic assessment policy because it shows how institutional conditions can affect teachers' assessment orientations. In addition, even though the majority of the teachers stated that they preferred and used alternative methods such

as self-assessment, peer-assessment, observation, and exit tickets, some MET stated that they preferred regular exams instead. They said that traditional exams were easy to implement and interpret. This finding aligns with Coombs et al. (2018), Mutlu (2025) and Chen (2021). Coombs et al. (2018) also reported that teachers who were in the early years of their careers showed a tendency to utilize “summative and standardized” practices (p. 142). This preference in the current research context can be explained in two ways. First, the institutional practices emphasize validity and reliability, which are provided by standardized exams. Therefore, rather than taking accountability, the MET may think that it is safer to implement these exams and get the results desired by the institution. Second, the majority of the participants are non-ELT graduates, and they do not get any testing and assessment courses during their undergraduate studies. They take one course on assessment after they graduate in their teacher certificate programs, but these courses are not sufficient in preparing them for the requirements of the job in terms of assessment. Also, no matter how many courses related to assessment are offered in teacher education programs, the non-ELT graduates cannot benefit from them since these courses are not offered within their degrees. In addition, while the HET might have developed their personal assessment philosophy and gained experience in using various formative assessment tools even though they did not get sufficient training during their undergraduate year and teaching certificate program, the MET might still be relying on traditional exams for clarity and efficiency due to lack of training and experience. This finding aligns with Çelebi and Kuşçuçuran’s (2019) study in which they talked about how undergraduate degree affected the teachers’ assessment preferences and how non-ELT graduates are more likely to implement more traditional assessment tools.

The teachers who preferred alternative assessment methods described them as essential and fun components of assessment. Yet, some of the HET were concerned about using portfolios and projects because they did not have enough experience and knowledge. One MET was worried about students using AI for their portfolios and projects. This may be due to personal experience or a lack of institutional regulations regarding the use of AI or plagiarism. The concern for AI use in alternative assessment tools is an important finding that indicates the teachers are already aware of the emerging challenges to assessment integrity.

Participants highlighted several challenges that hinder their ability to design, implement, and interpret effective assessment practices. Time constraints were frequently mentioned by the participants. Almost all participants reported that they struggled to find enough time to effectively plan, implement and provide feedback to assessments. This may reduce the effectiveness of assessments and the quality of feedback, which threatens the main aim of formative assessment practices. This finding might have stemmed from the heavy workload associated with standardized summative assessment practices such as blind double grading, standardization procedures and spot checking and their lesson planning responsibilities. These demands leave teachers with limited time and energy to plan formative assessment activities that can provide rich feedback about student learning and progress. Similar concerns and issues have been observed in the Turkish context where teachers reported not having enough time to design effective assessment tools and being pushed towards summative assessments for efficiency and practicality (Kaya & Mede, 2021; Çimen, 2022).

The second challenge was institutional constraints, as previously stated, the lack of autonomy in designing summative assessments conflicts with equitable and differentiated fairness approaches. In this specific research context, this finding might be due to the role of the testing unit and the little to no autonomy that teachers have over summative assessment tools. It shows parallelism with many studies in the Turkish context such as Yeni-Palabıyık and Daloğlu (2025), Öz and Atay (2017) and Mutlu (2025), who also stated that despite the desire to implement formative assessment tools and modify summative assessments, the teachers are constrained by structural constraints and institutional demands. Therefore, both the present study and the previous studies demonstrate how teachers' beliefs and personal assessment approaches are outweighed by institutional culture in the Turkish context.

The final challenge was put forward only by MET. It was about the negative impact of assessment on teacher well-being. The teachers expressed the emotional burden of having to work for long periods to grade assessments effectively and the feeling of not doing enough. One possible explanation why this was only mentioned by the MET might be because emotional labor and burnout have been more emphasized in

recent teacher education programs, so the MET may be more likely to articulate the emotional challenges of the job.

Lastly, a strong need for professional development activities was expressed during the interviews. Participants highlighted the need for collaborative, practical, and sustainable professional development opportunities. Many similar studies reported a similar need for collaborative, interactive and sustainable professional development activities (Ashraf & Zolfaghar, 2015; Jannati,2015; DeLuca et al., 2018; Mertler, 2003; Xu & Brown, 2016; Büyükkarcı, 2016). This consistent desire for additional PD activities might indicate that the mandatory training selected by the head of the department might not be perceived as sufficient by the teachers and they might be tailored to individual teacher needs.

CHAPTER 6

CONCLUSION

The chapter presents the key findings, concluding remarks, the implications and limitations of the study.

6.1 Conclusion

The current study aimed to examine the assessment approaches of in-service EFL teachers working at a private middle school in Ankara in relation to four themes of assessment specified within the ACAI framework and to explore whether teaching experience played a role in their assessment approaches. The findings revealed that teachers supported the assessment for learning approach as they viewed assessment as a tool to improve teaching and enhance student learning. However, summative assessments dominated their teaching activities due to institutional factors such as the number of summative assessments implemented in a semester or time and pacing concerns. Therefore, institutional constraints overshadowed teacher beliefs and experiences as the teachers could not implement formative assessment practices to the extent that they desired due to institutional demands. Furthermore, the teachers emphasized the importance of designing reliable assessments. However, the study revealed a gap between their beliefs regarding fairness and their applications because they were not given the opportunity to design their own summative assessments or modify or adapt the existing ones. The negative impact of assessment on teacher well-being appeared as a dimension that was directly influenced by assessment practices. Finally, teaching experience led to some differences between the experience groups as teachers with more experience were more confident in using alternative assessment methods, and some members of the group with less experience tended to rely on traditional assessment tools for clarity and efficiency.

It is hoped that the study offers important insights into assessment research by providing context specific information. Even though it is a case study, the current research context is considered to be a good example of private school contexts because it is a large private institution with a centralized testing unit and top-down summative assessment policy, which is the case for the majority of private institutions in Türkiye as it was stated by Yıldırım-Seheryeli and Gelbal (2022). Moreover, it highlights the gap between teachers' assessment beliefs and practices, which aligns with previous studies in the Turkish context by showing how systemic factors can restrict both teachers with high and low levels of teaching experience. Finally, the study draws attention to the fact that non-ELT graduates may rely more on traditional assessment since they do not receive any assessment training during their undergraduate degree.

6.2. Implications of the study

Based on these results, this teacher-oriented case study revealed important suggestions for teacher educators, professional development program developers and school administrators.

MET demonstrated stronger support for ACAI principles compared to their more experienced colleagues. The MET group in the current study showed stronger support for ACAI principles in almost every theme. However, they also reported uncertainty and a lack of confidence in implementing assessment tools. Therefore, it can be said that classroom assessment approaches should be more explicitly integrated into teacher education programs.

They should include courses which include hands-on practice in designing, using and adapting classroom assessments. The initial teacher education programs should aim to equip pre-service teachers not only with theoretical knowledge but also with experience and confidence in assessment practices. Also, as it was reported by Hatipoğlu (2015), pre-service teachers only get one course on assessment, and it is not sufficient to prepare teacher candidates for the requirements of classroom assessment. Therefore, multiple courses that focus on not only the technical aspect of

assessment but also on alternative assessment, fairness and ethics in assessment, and language-specific assessment approaches in the teaching of English as a foreign language should be added to the curriculum. Furthermore, special attention should be devoted to non-ELT graduates who often do not get formal and sufficient assessment training before they get into the profession. As Çelebi and Kuşçuran (2019) explained, these teachers might show heavier reliance on traditional assessment tools. Therefore, PD programs and teacher certificate programs should provide tailored and targeted support for non-ELT graduates to familiarize them with alternative and formative assessment tools and help them gain confidence by providing practical applications of these tools.

It was evident that developing the teachers' understanding and implementation of fairness in assessment practices was an urgent need when the participants' endorsement of the approaches within the assessment fairness theme was analyzed. Several teachers associated fairness with consistency and equality in scoring through processes such as double grading and standardization, and they expressed an urgent need for practical, collaborative and long-term PD opportunities. Therefore, PD programs should help teachers reconceptualize fairness in assessment, highlight the importance of differentiated fairness practices and provide the teachers with meaningful implementations of differentiated assessment through hands-on and interactive activities.

Teachers continuously reported that they had a lack of autonomy in summative assessment, and this lack of autonomy hindered the summative assessment practices and prevented teachers from making assessments fair for all of their students. School administrators should reconsider these top-down assessment practices and the effect of standardized tests and allocate more space and opportunities for teacher input and autonomy in the process of designing summative assessments. Furthermore, MET expressed the emotional burden of the assessment practices. They argued that they had to spend a lot of time grading, planning and implementing assessments during some periods, which had a negative impact on their well-being. Therefore, institutions should recognize and address teacher workload and burnout and reform their policies by taking this into account.

Finally, the study had important implications for policy makers as well. As explained by Yeni-Palabıyık and Dalođlu (2025), the ministry of national education has revised the assessment policies to make them more student centered and formative oriented. However, the gap between the actual practices and the policy requirements has also been emphasized by the current study. Therefore, education authorities should make some efforts to bridge the gap between teachers' perceptions and practices and curriculum requirements. The exam driven nature of the Turkish educational system leads to limited use of formative and alternative assessment tools and creates conflict between teachers' beliefs and desires and institutional requirements. The policy makers should keep the teacher education standards updated, provide PD focusing on assessment and create more space for contemporary assessment approaches and practices.

6.3. Limitations of the study and recommendations for further research

This case study was conducted with a small sample within a specific context. Therefore, the results may not be generalized to larger populations, different educational contexts and undergraduate backgrounds. For future studies, the study can be conducted with a larger sample from various educational contexts and undergraduate backgrounds to compare the assessment orientations of ELT and non-ELT graduates.

Also, the study relied on self-reported data from the questionnaire and semi-structured interviews. The participants may have given their answers based on pedagogical appropriateness rather than actual teaching practices. In future studies, classroom observations can also be used to compare the participants' self-reported assessment practices with actual teaching practices.

Lastly, the study could not explore how the teachers would design their summative assessments if they had been given the chance and opportunity because they had limited autonomy in designing them. Future studies can investigate how teachers would design their summative assessment tools when given the chance or provide them with opportunities to design their own summative assessment tools to better understand their priorities and values.

REFERENCES

- American Federation of Teachers, National Council on Measurement in Education, & National Education Association. (1990). Standards for teacher competence in educational assessment of students. *Educational Measurement: Issues and Practice*, 9(4), 30–32. <https://doi.org/10.1111/j.1745-3992.1990.tb00391.x>
- Amirian, S. M. R. (2025). Understanding EFL teachers' formative assessment literacy: Insights from a mixed-methods study. *Language Testing in Asia*, 15(25). <https://doi.org/10.1186/s40468-025-00363-y>
- Ballıdağ, S., & Inan Karagül, B. (2021). Exploring the language assessment literacy of Turkish in-service EFL teachers. *Bahkesir Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 24(45), 73–92. <https://doi.org/10.31795/baunsobed.909953>
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74. <https://doi.org/10.1080/0969595980050102>
- Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice*, 25(6), 551–575. <https://doi.org/10.1080/0969594X.2018.1441807>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Brookhart, S. M. (2001). Successful students' formative and summative uses of assessment information. *Assessment in Education: Principles, Policy & Practice*, 8(2), 153–169. <https://doi.org/10.1080/09695940123775>

- Büyükkarıcı, K. (2016). Identifying the areas for English language teacher development: A study of assessment literacy. *Pegem Journal of Education and Instruction*, 6(3), 333–346. <https://doi.org/10.14527/pegegog.2016.017>
- Campbell, S., Greenwood, M., Prior, S., Shearer, T., Walkem, K., Young, S., Bywaters, D., & Walker, K. (2020). Purposive sampling: Complex or simple? Research case examples. *Journal of Research in Nursing*, 25(8), 652–661. <https://doi.org/10.1177/1744987120927206>
- Çayır, A. (2022). *Pre-service science teachers' approaches to classroom assessment* (Master's thesis). Middle East Technical University.
- Chang, D. Y. S., Lin, M. H., & Lee, J. Y. (2024). Exploring language assessment literacy and needs of English teachers at senior high school level. *Asia Pacific Journal of Education*, 44(4), 854–872. <https://doi.org/10.1080/02188791.2024.2313486>
- Chappuis, S., & Stiggins, R. J. (2002). Classroom assessment for learning. *Educational Leadership*, 60(1), 40–44.
- Chen, Y. (2021). Exploring secondary school EFL teachers' assessment literacy in practice: A case study in China. *English Language Teaching*, 14(12), 1–7. <https://doi.org/10.5539/elt.v14n12p1>
- Çimen, S. S. (2022). Exploring EFL assessment in Turkey: Curriculum and teacher practices. *International Online Journal of Education and Teaching*, 9(1), 531–550.
- Coombe, C., Vafadar, H., & Mohebbi, H. (2020). Language assessment literacy: What do we need to learn, unlearn, and relearn? *Language Testing in Asia*, 10(1), Article 3. <https://doi.org/10.1186/s40468-020-00101-6>

- Coombs, A. J. (2017). *Teacher educators' approaches to assessment* (Master's thesis, Queen's University, Canada). Queen's University Research Repository.
- Coombs, A., DeLuca, C., LaPointe-McEwan, D., & Chalas, A. (2018). Changing approaches to classroom assessment: An empirical study across teacher career stages. *Teaching and Teacher Education*, *71*, 134–144. <https://doi.org/10.1016/j.tate.2017.12.010>
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.
- Danielson, C. (2013). *The framework for teaching evaluation instrument*. Danielson Group.
- Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing*, *25*(3), 327–347. <https://doi.org/10.1177/0265532208090156>
- DeLuca, C., & Klinger, D. A. (2010). Assessment literacy development: Identifying gaps in teacher candidates' learning. *Assessment in Education: Principles, Policy & Practice*, *17*(4), 419–438. <https://doi.org/10.1080/0969594X.2010.516643>
- DeLuca, C., Coombs, A., & LaPointe-McEwan, D. (2019). Assessment mindset: Exploring the relationship between teacher mindset and approaches to classroom assessment. *Studies in Educational Evaluation*, *61*, 159–169. <https://doi.org/10.1016/j.stueduc.2019.03.012>
- DeLuca, C., Coombs, A., MacGregor, S., & Rasooli, A. (2019). Toward a differential and situated view of assessment literacy: Studying teachers' responses to classroom assessment scenarios. *Frontiers in Education*, *4*, Article 94. <https://doi.org/10.3389/educ.2019.00094>

- DeLuca, C., LaPointe-McEwan, D., & Luhanga, U. (2016a). Approaches to Classroom Assessment Inventory: A new instrument to support teacher assessment literacy. *Educational Assessment*, 21(4), 248–266. <https://doi.org/10.1080/10627197.2016.1236677>
- DeLuca, C., LaPointe-McEwan, D., & Luhanga, U. (2016b). Teacher assessment literacy: A review of international standards and measures. *Educational Assessment, Evaluation and Accountability*, 28(3), 251–272. <https://doi.org/10.1007/s11092-015-9233-6>
- DeLuca, C., Valiquette, A., Coombs, A., LaPointe-McEwan, D., & Luhanga, U. (2018). Teachers' approaches to classroom assessment: A large-scale survey. *Assessment in Education: Principles, Policy & Practice*, 25(4), 355–375. <https://doi.org/10.1080/0969594X.2016.1244514>
- DiCarlo, K., & Cooper, L. (2014). Classroom assessment techniques: A literature review. *Journal of Instructional Research*, 3, 15–20.
- Doğan, Ö., & Ünal, B. (2024). Refining the assessment literacy competence among K12 EFL teachers. *Participatory Educational Research*, 11(4), 250–266. <https://doi.org/10.17275/per.24.59.11.4>
- Fındıklı, S., & Büyükkarcı, K. (2024). A comparison of Turkish and European English language teachers' language assessment knowledge levels and perceptions. *International Journal of Assessment Tools in Education*, 11(2), 345–367. <https://doi.org/10.21449/ijate.1360899>
- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113–132. <https://doi.org/10.1080/15434303.2011.642041>
- Ghaicha, A. (2016). Theoretical framework for educational assessment: A synoptic review. *Journal of Education and Practice*, 7(24), 212–231.

- Hakim, B. (2015). English language teachers' ideology of ELT assessment literacy. *International Journal of Education and Literacy Studies*, 3(4), 42–48. <https://doi.org/10.7575/aiac.ijels.v.3n.4p.42>
- Hatipoğlu, Ç. (2015). English language testing and evaluation, ELTE training in Turkey: Expectations and needs of pre-service English language teachers. *ELT Journal*, 69(2), 111–128. <https://hdl.handle.net/11511/78435>
- Işık, A., & Sarı, R. (2021). Are English language teachers assessment literate? *Çukurova Üniversitesi Eğitim Fakültesi Dergisi*, 50(2), 907–928. <https://doi.org/10.14812/cuefd.877706>
- Jannati, S. (2015). ELT teachers' language assessment literacy: Perceptions and practices. *The International Journal of Research in Teacher Education*, 6(2), 26–37.
- Joint Committee on Standards for Educational Evaluation. (2015). *Classroom assessment standards: Practices for PK-12 teachers*. <http://www.jcsee.org/the-classroom-assessment-standards-new-standards>
- 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*, 6(1), 223–232.
- Kaya, T., & Mede, E. (2021). Exploring language assessment literacy of EFL instructors in language preparatory programs. *İstanbul Aydın Üniversitesi Eğitim Fakültesi Dergisi*, 7(1), 163–189.
- Kelly, S. E. (2010). Qualitative interviewing techniques and styles. In I. Bourgeault, R. Dingwall, & R. de Vries (Eds.), *The SAGE handbook of qualitative methods in health research* (pp. 307–326). SAGE Publications. <https://doi.org/10.4135/9781446268247.n17>

- Kothai, N., & Rajendran, M. (2020). Pre-service teachers' approaches to classroom assessment. *Humanities & Social Sciences Reviews*, 8(1), 666–673. <https://doi.org/10.18510/hssr.2020.8180>
- Kremmel, B., & Harding, L. (2020). Towards a comprehensive, empirical model of language assessment literacy across stakeholder groups: Developing the Language Assessment Literacy Survey. *Language Assessment Quarterly*, 17(1), 100–120. <https://doi.org/10.1080/15434303.2019.1674855>
- Latif, M. W., & Wasim, A. (2024). Investigating EFL instructors' approaches to classroom-based assessment culture: An explanatory sequential mixed-method approach. *Language Testing in Asia*, 14(57). <https://doi.org/10.1186/s40468-024-00332-x>
- Lazarakou, E. (2025). Greek pre-service teachers' approaches and beliefs regarding classroom assessment. *European Journal of Education and Pedagogy*, 6(1), 83–89. <https://doi.org/10.24018/ejedu.2025.6.1.922>
- Lincoln, Y. S., Guba, E. G., & Pilotta, J. J. (1985). Naturalistic inquiry. *International Journal of Intercultural Relations*, 9(4), 438–439. [https://doi.org/10.1016/0147-1767\(85\)90062-8](https://doi.org/10.1016/0147-1767(85)90062-8)
- Lynch, B. K. (2001). Rethinking assessment from a critical perspective. *Language Testing*, 18(4), 351–372. <https://doi.org/10.1177/026553220101800403>
- Malone, M. E. (2013). The essentials of assessment literacy: Contrasts between testers and users. *Language Testing*, 30(3), 329–344. <https://doi.org/10.1177/0265532213480129>
- Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation*. Jossey-Bass.

- Mertler, C. A. (2003). Secondary teachers' assessment literacy: Does classroom experience make a difference? *American Secondary Education*, 33(1), 49–64.
- Mutlu, A. (2025). Turkish EFL teachers' beliefs, practices and challenges on assessment. *Manisa Celal Bayar University International Journal of English Language Studies*, 4(1), 1–19.
- Ölmezer Öztürk, E. (2021). Developing language assessment literacy of EFL pre-service teachers through classroom assessment course. *International Journal of Contemporary Educational Research*, 8(3), 13–22. <https://doi.org/10.33200/ijcer.932721>
- Önalın, O. K., & Esim, G. (2020). Primary and secondary level EFL teachers' use of assessment and assessment results in Turkey. *i-manager's Journal on English Language Teaching*, 10(3), 1–11. <https://doi.org/10.26634/jelt.10.3.16387>
- Opre, D. (2015). Teachers' conceptions of assessment. *Procedia – Social and Behavioral Sciences*, 209, 229–233. <https://doi.org/10.1016/j.sbspro.2015.11.222>
- Öz, S., & Atay, D. (2017). Turkish EFL instructors' in-class language assessment literacy: Perceptions and practices. *ELT Research Journal*, 6(1), 25–44.
- Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015). Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Administration and Policy in Mental Health*, 42(5), 533–544. <https://doi.org/10.1007/s10488-013-0528-y>
- Pastore, S. (2023). Teacher assessment literacy: A systematic review. *Frontiers in Education*, 8, Article 1217167. <https://doi.org/10.3389/educ.2023.1217167>

- Phongsirikul, M. (2018). Traditional and alternative assessments in ELT: Students' and teachers' perceptions. *REFlections*, 25(1), 61–84.
- Phye, G. D. (1996). *Handbook of classroom assessment: Learning, achievement, and adjustment*. Academic Press.
- Pill, J., & Harding, L. (2013). Defining the language assessment literacy gap: Evidence from a parliamentary inquiry. *Language Testing*, 30(3), 381–402. <https://doi.org/10.1177/0265532213480337>
- 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. <https://doi.org/10.1111/j.1745-3992.1993.tb00548.x>
- Price, J. K., Light, D., & Pierson, E. (2014). Classroom assessment: A key component to support education transformation. In D. Light & E. Pierson (Eds.), *ICT in education in global context: Emerging trends report 2013–2014* (pp. 31–46). Springer. https://doi.org/10.1007/978-3-662-43927-2_3
- Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory Into Practice*, 48(1), 4–11. <https://doi.org/10.1080/00405840802577536>
- Popham, W. J. (2024). *Classroom assessment: What teachers need to know* (10th ed.). Pearson.
- Riconscente, M. M., Mislevy, R. J., & Hamel, L. (2005). *An introduction to PADI task templates* (PADI Technical Report No. 3). SRI International.
- Roever, C., & McNamara, T. (2006). Language testing: The social dimension. *International Journal of Applied Linguistics*, 16(2), 242–258.

- Sarıyıldız Canlı, G., & Altay, İ. F. (2023). Training needs of in-service EFL teachers in language testing and assessment. *Pegem Journal of Education and Instruction*, 14(1), 69–79. <https://doi.org/10.47750/pegegog.14.01.08>
- Shepard, L. A. (2019). Classroom assessment to support teaching and learning. *The ANNALS of the American Academy of Political and Social Science*, 683(1), 183–200. <https://doi.org/10.1177/0002716219843816>
- Soodmand Afshar, H., & Ranjbar, N. (2021). EAP teachers' assessment literacy: From theory to practice. *Studies in Educational Evaluation*, 70, 101042. <https://doi.org/10.1016/j.stueduc.2021.101042>
- Stiggins, R. J. (1991). Front matter. *The Phi Delta Kappan*, 72(7), 534–539. <https://www.jstor.org/stable/20404446>
- Sultana, N. (2019). Language assessment literacy: An uncharted area for the English language teachers in Bangladesh. *Language Testing in Asia*, 9(1), Article 1. <https://doi.org/10.1186/s40468-019-0077-8>
- Taylor, L. (2013). Communicating the theory, practice and principles of language testing to test stakeholders: Some reflections. *Language Testing*, 30(3), 403–412. <https://doi.org/10.1177/0265532213480338>
- Tierney, R. D. (2006). Changing practices: Influences on classroom assessment. *Assessment in Education: Principles, Policy & Practice*, 13(3), 239–264. <https://doi.org/10.1080/09695940601035387>
- Tosuncuoğlu, I. (2018). Importance of assessment in ELT. *Journal of Education and Training Studies*, 6(9), 163–167.
- Tuncer, M., & Merc, A. (2023). Turkish high school EFL teachers' AfL practices: Reported vs. actual classroom practices. *Journal of Innovative Research in Teacher Education*, 4(2), 201–221. <https://doi.org/10.29329/jirte.2023.572.3>

- Vogt, K., & Tsagari, D. (2014). Assessment literacy of foreign language teachers: Findings of a European study. *Language Assessment Quarterly*, 11(4), 374–402. <https://doi.org/10.1080/15434303.2014.960046>
- Volante, L., & Fazio, X. (2007). Exploring teacher candidates' assessment literacy: Implications for teacher education reform and professional development. *Canadian Journal of Education*, 30(3), 749–770. <https://doi.org/10.2307/20466661>
- Webb, N. L. (2002, April). Assessment literacy in a standards-based urban education setting. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA, United States.
- Weng, F., & Shen, B. (2022). Language assessment literacy of teachers. *Frontiers in Psychology*, 13, 864582. <https://doi.org/10.3389/fpsyg.2022.864582>
- Willis, J., Adie, L., & Klenowski, V. (2013). Conceptualising teachers' assessment literacies in an era of curriculum and assessment reform. *Australian Educational Researcher*, 40(2), 241–256. <https://doi.org/10.1007/s13384-013-0089-9>
- Xu, Y., & Brown, G. T. L. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, 58, 149–162. <https://doi.org/10.1016/j.tate.2016.05.010>
- Xu, Y., & Brown, G. T. L. (2017). University English teacher assessment literacy: A survey-test report from China. *Papers in Language Testing and Assessment*, 6(1), 133–158. <https://doi.org/10.58379/UZON5145>
- Yastibaş, A. E., & Takkaç, M. (2018). Understanding language assessment literacy: Developing language assessment. *Journal of Language and Linguistic Studies*, 14(1), 178–193.

- Yeni-Palabıyık, P., & Dalođlu, A. (2025). Policy and practice in L2 classroom assessment: Policy implementation at a state high school in Türkiye. *Education Inquiry*. Advance online publication. <https://doi.org/10.1080/20004508.2025.2453278>
- Yıldırım Seheryeli, M., & Gelbal, S. (2020). Practices and opinions of teachers working at public, private and International Baccalaureate schools on measurement and evaluation [Devlet, özel ve Uluslararası Bakalorya okullarındaki öğretmenlerin ölçme ve değerlendirme çalışmaları ve görüşleri]. *International Journal of Curriculum and Instructional Studies*, 10(1), 221–260.
- Yin, R. K. (2014). *Case study research: Design and methods* (5th ed.). SAGE Publications.
- Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). SAGE Publications.
- Zhang, Z., & Burry-Stock, J. A. (1997, March). Assessment practices inventory: A multivariate analysis of teachers' perceived assessment competency (ERIC Document No. ED409311). Paper presented at the annual meeting of the American Educational Research Association (AERA), Chicago, IL. ERIC. <https://eric.ed.gov/?id=ED409311>
- Zolfaghari, S., & Ashraf, H. (2015). The relationship between EFL teachers' assessment literacy, their teaching experience, and their age: A case of Iranian EFL teachers. *Theory and Practice in Language Studies*, 5(12), 2550–2556. <https://doi.org/10.17507/tpls.0512.16>

APPENDICES

A. APPROVAL OF THE METU HUMAN SUBJECTS ETHICS COMMITTEE

UYGULAMALI ETİK ARAŞTIRMA MERKEZİ
APPLIED ETHICS RESEARCH CENTER

DUMLUPINAR BULVARI 06800
ÇANKAYA ANKARA/TURKEY
T: +90 312 210 22 91
F: +90 312 210 79 59
ueam@metu.edu.tr
www.ueam.metu.edu.tr



ORTA DOĞU TEKNİK ÜNİVERSİTESİ
MIDDLE EAST TECHNICAL UNIVERSITY

Konu : Değerlendirme Sonucu 30.06.2025
Gönderen : ODTÜ İnsan Araştırmaları Etik Kurulu
İlgi : İnsan Araştırmaları Etik Başvurunuz

Sayın Gözdenur Çetin,
Sayın Dr. Öğr. Üyesi Müge Gündüz,

“Investigating the Assessment Approaches of EFL Teachers in a Private School in Ankara: Impact of Teaching Experience” başlıklı araştırmanız ODTÜ İnsan Araştırmaları Etik Kurulu tarafından uygun görülerek 0327-ODTÜIAEK-2025 protokol numarası ile onaylanmıştır.

Bilgilerinize sunarım.

Prof. Dr. Ş. Halil TURAN
Başkan

Prof. Dr. İ. Semih AKÇOMAK
Üye

Doç. Dr. Ali Emre TURGUT
Üye

Doç. Dr. Aşlı KILIÇ ÖZHAN
Üye

Doç. Dr. Çağrı TOPAL
Üye

Doç. Dr. Pınar AYKAÇ LEIDHOLM
Üye

Dr. Öğr. Üyesi Müge GÜNDÜZ
Üye

B. ADAPTED VERSION OF THE ACAI

PART A: PARTICIPANT BACKGROUND INFORMATION

Please answer the following questions.

- 1. Which gender do you self-identify with?**
 - a. Male
 - b. Female
 - c. Prefer not to respond

- 2. How old are you?: _____**

- 3. Which university and department did you graduate from?**

- 4. How long have you been teaching?**
_____.

- 5. How long have you been teaching at this school?**
_____.

- 6. What grade level(s) are you currently teaching?**
_____.

- 7. Have you attended any professional development workshops, seminars, or trainings on assessment in the past 3 years?**
 - a. Yes
 - b. No

If yes, please specify:

- 8. How confident do you feel in your ability to design effective classroom assessments?**
 - a) Not confident at all
 - b) Slightly confident
 - c) Somewhat confident
 - d) Very confident
 - e) Extremely confident

- 9. How knowledgeable do you consider yourself about classroom assessment principles and practices?**
 - a. Not knowledgeable at all
 - b. Slightly knowledgeable
 - c. Somewhat knowledgeable
 - d. Very knowledgeable
 - e. Extremely knowledgeable

PART B: SCENARIO BASED QUESTIONS**Scenario 1:**

You give your class a paper-pencil summative unit test with accommodation and modifications for identified learners. Ten of the 24 students fail.

As a teacher in this situation, how likely are you to do each of the following actions as part of your overall response to the scenario? If you have experienced this scenario in your teaching context, base your response on the actions you took.

Action	Highly Unlikely (1)	Unlikely (2)	Somewhat Unlikely (3)	Somewhat Likely (4)	Likely (5)	Highly Likely (6)	Don't Know
Keep and record the test grade for all students and move on to the next topic.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Disregard the test grade, then reteach and retest the material.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Have students generate a plan to relearn the material.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Analyze test results to determine if failure was the result of poor test design.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Remove test questions that most students struggled with and recalculate student scores without those questions.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Schedule student conferences to discuss test performance and next steps.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Allow the entire class to retake a similar test.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Allow only students with exceptionalities who struggled with the test to take a new assessment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Provide a new assessment to every student who struggled with the test.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Identify test questions that most students consistently answered incorrectly, then provide students with new questions to re-test those concepts.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Adjust student grades based on previously collected formative assessment information.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Analyze test results to identify test design errors and consider student circumstances for failure, then adjust individual student grades accordingly.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

If your response to this scenario was not listed above, how would you most likely respond?

Scenario 2:

You discover that one of your students has plagiarized sections of their term project.

As a teacher in this situation, how likely are you to do each of the following actions? If you have experienced this scenario in your teaching context, base your response on the actions you took.

Action	Highly Unlikely (1)	Unlikely (2)	Somewhat Unlikely (3)	Somewhat Likely (4)	Likely (5)	Highly Likely (6)	Don't Know
Reduce the student's grade.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Have them re-write the plagiarized section in their own words, then re-grade the assignment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ask them to reflect on why plagiarism is a problem and what they would do differently next time.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
As the teacher, reflect on how the assignment could have been structured differently to deter plagiarism.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Adjust the student's grade to reflect the portion of work that was plagiarized.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Discuss with the student the reasons for the plagiarism, severity of plagiarism, and negotiate potential next steps for their learning.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Apply the same consequence you would for other students to ensure all students are treated equally.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Consider if the student has identified learning needs and exceptionalities before determining a response to plagiarism.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Discuss why the student plagiarized and agree upon an appropriate alternative assignment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Apply all aspects of school policy on plagiarism to ensure consistency across all students.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Consider the original aspects of the assignment and the plagiarized text to determine what the student knows and does not appear to know related to content expectations.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Apply the school policy on plagiarism unless there is a good reason that explains why the student plagiarized.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

If your response to this scenario was not listed above, how would you most likely respond?

Scenario 3:

Out of 28 students in your class, 4 students are classified/identified with an exceptionality and have an Individual Education Plan (IEP) (i.e., each student requires accommodations, but not a modified curriculum). Additionally, several other students have different learning needs that are not officially identified. You must decide how to accurately assess learning in your class.

As a teacher in this situation, how likely are you to do each of the following actions? If you have experienced this scenario in your teaching context, base your response on the actions you took.

Action	Highly Unlikely (1)	Unlikely (2)	Somewhat Unlikely (3)	Somewhat Likely (4)	Likely (5)	Highly Likely (6)	Don't Know
Provide all identified students with accommodations on all summative assessments.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Use formative assessments with all of your students based on their individual learning needs to support their learning.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Allow each student to develop a learning plan based on their strengths and goals to support their own learning.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Design a variety of assessment methods to target each learning objective.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Develop and use different rubric criteria to assess identified students.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Explain to students and parents the purpose of accommodations and how they will be implemented to support student learning and assessment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Grade students using the same assessments and without accommodations.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Assume that accommodated assessments provide an equally valid measure of identified students' learning.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Provide a variety of assessment options and allow students to choose how they will demonstrate their learning.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Use the same scoring rubric for all students.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Consider individual student learning needs when assessing student work.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Use the same scoring rubric for all students but apply professional judgment to adapt criteria based on individual student needs.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

If your response to this scenario was not listed above, how would you most likely respond?

Scenario 4:

"You are preparing a new unit for your English class. As part of your planning, you must decide how to design assessments, how to use them during the unit, and how to evaluate student learning at the end."

As a teacher in this situation, how likely are you to do each of the following actions? If you have experienced this scenario in your teaching context, base your response on the actions you took.

Action	Highly Unlikely (1)	Unlikely (2)	Somewhat Unlikely (3)	Somewhat Likely (4)	Likely (5)	Highly Likely (6)	Don't Know
Start by designing a summative assessment based on curriculum expectations, then plan your lessons.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Use formative assessments to guide your subsequent lesson planning.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Co-construct learning goals with students, then have each student develop a personal learning plan for the unit.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Design a summative unit assessment based on what was taught to students.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
When planning a unit, consider when assessments and feedback will be provided to students.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Discuss assignments and co-construct grading criteria for the unit with your students.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Plan class lessons and assessments that are the same for all students and encompass the curriculum expectations.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ensure that assessments throughout the unit are accommodated for formally identified learners (i.e., learners with exceptionalities).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Give all students a diagnostic assessment at the beginning of the unit and have students use their results to select personalized appropriate learning and assessment activities.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Work with other teachers to develop a reliable summative assessment for the unit.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
In designing your assessments, prioritize the alignment between the assessment and the intended learning objectives.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Adapt assessments that have worked well with other students like yours to take into consideration the context and activities of your enacted lessons.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

If your response to this scenario was not listed above, how would you most likely respond?

C. INTERVIEW QUESTIONS IN ENGLISH

1. How long have you been teaching English, and what grade levels do you currently teach?
2. Have you taken any professional development courses on classroom assessment?
If yes, do you feel they adequately prepared you for real classroom assessment? Why or why not?
3. What do you think is the main purpose of classroom assessment in ELT?
4. Which assessment methods do you prefer? (e.g., tests, projects, portfolios, self-assessment, peer assessment, oral exams) Why?
5. Do you feel confident using alternative assessments (e.g., portfolios, peer-assessment, self-assessment)? Why or why not?
6. What challenges do you face while implementing assessments in your classroom?
7. In your own words, what does 'fair assessment' mean to you, and how do you try to ensure fairness in your classroom?
8. Have you ever faced challenges in making assessment fair for all students? If so, how did you address them?
9. How would you describe your personal approach to assessment?
10. Have you ever felt conflicted between being a facilitator of learning and an evaluator? How do you handle this?
11. Do you think your level of experience influences the way you assess students in any way? If so, how?
12. What kind of training or professional development would help you feel more confident in assessing students?
13. What are your thoughts on your institution's assessment policy? What do you think are its strengths and areas for improvement?

D. INTERVIEW QUESTIONS IN TURKISH

1. Ne kadar süredir İngilizce öğretiyorsunuz ve şu anda hangi sınıf seviyelerine ders veriyorsunuz?
2. Daha önce sınıf içi ölçME ve değerlendirme üzerine herhangi bir hizmet içi eğitim veya seminer aldınız mı?

Eğer aldıysanız, bu eğitimlerin sizi gerçek sınıf uygulamalarına hazırladığını düşünüyor musunuz? Neden?
3. Sizce İngilizce öğretimde sınıf içi değerlendirmenin temel amacı nedir?
4. Hangi ölçme ve değerlendirme yöntemlerini tercih ediyorsunuz? (örneğin testler, projeler, portfolyolar, özdeğerlendirme, akran değerlendirmesi, sözlü sınavlar) Neden?
5. Alternatif değerlendirme yöntemlerini, (örneğin portfolyo, akran değerlendirmesi, özdeğerlendirme) kullanma konusunda kendinizi ne kadar yetkin hissediyorsunuz? Neden?
6. Sınıfınızda ölçme ve değerlendirme yaparken karşılaştığınız zorluklar nelerdir?
7. Sizin için adil değerlendirme ne anlama geliyor? Sınıfınızda adaleti sağlamak için neler yapıyorsunuz?
8. Tüm öğrenciler için adil bir değerlendirme yapmakta zorlandığınız durumlar oldu mu? Olduysa bu durumu nasıl yönettiniz?
9. Kendi değerlendirme yaklaşımınızı nasıl tanımlarsınız?
10. Hiç öğrenmeyi kolaylaştıran bir öğretmen olma ile öğrencileri değerlendiren biri olma rolleriniz arasında çelişki yaşadınız oldu mu? Bu durumu nasıl aşılıyorsunuz?
11. Sizce öğretmenlikteki deneyim seviyeniz öğrenci değerlendirme şeklinizi etkiliyor mu? Etkiliyorsa nasıl?
12. Öğrencileri değerlendirme konusunda daha donanımlı hissetmenizi sağlayacak ne tür bir eğitim veya hizmet içi gelişim programı size yardımcı olurdu?
13. Kurumunuzun mevcut ölçme ve değerlendirme politikası hakkında ne düşünüyorsunuz? Sizce güçlü yönleri ve geliştirilebilecek yönleri nelerdir?

E. INFORMED CONSENT FORM

INFORMED CONSENT FORM

You are being invited to participate in a research study conducted as part of a master's thesis by Gözdenur Çetin, under the supervision of Asst. Prof. Dr. Prof. Dr. Müge Gündüz, in the Department of Foreign Language Education at Middle East Technical University (METU). Please read the following information carefully, and feel free to ask any questions to the researcher if anything is unclear or if you would like more information.

Purpose of the Study:

The purpose of this study is to examine how in-service EFL teachers approach classroom assessment in a private middle school setting, and to explore how teaching experience influences their assessment approaches. The findings will contribute to a better understanding of teachers' assessment literacy and practices and will help inform future professional development and assessment policy in similar educational contexts.

Research Procedures:

If you agree to participate, you will be asked to complete a background information questionnaire, the Approaches to Classroom Assessment Inventory (ACAI) and participate in a semi-structured interview about your classroom assessment beliefs and practices.

The background questionnaire and ACAI will take approximately 30–40 minutes in total. The semi-structured interview will take approximately 45 minutes and will be conducted in person or online, depending on your preference. The interview will be audio recorded with your consent. All data will be securely stored on password-protected devices and used only for research purposes. Interview recordings will be transcribed for analysis.

Voluntary Participation:

Your participation in this study is completely voluntary. You may choose not to answer any question or to withdraw from the study at any time without giving a reason and without any consequences. If you choose to withdraw before data collection is completed, your data will be returned to you or securely destroyed.

Confidentiality:

All information collected in this study will be kept strictly confidential. Your identity

will not be revealed in any reports, publications, or presentations. Pseudonyms will be used in all notes, transcripts, and documents. The data will be stored securely.

Benefits and Risks:

There are no known risks associated with participating in this study. You may benefit from reflecting on your own assessment beliefs and practices. The study’s findings may contribute to enhancing professional development in assessment for EFL teachers in Türkiye.

Contact Information:

If you have any questions about this study or experience any concerns related to your participation, you may contact the researcher.

Researcher:

Name – Surname: Gözdenur Çetin

Department: METU Department of Foreign Language Education

Phone:

Email:

I have read and understood the information provided above and have had the opportunity to ask questions. I understand that my participation is voluntary, and I may withdraw at any time without giving a reason. I understand that I will receive a copy of this consent form. I voluntarily agree to participate in this study.

Participant’s Name: _____

Participant’s Signature: _____

Date: _____

F. THE CODEBOOK

Parent Code	Child Code	Description
Assessment Purpose and Use	Using assessment to understand needs and track progress	Assessment is used to identify student needs and record students' progress
	Using assessment as a learning tool	Assessment is used to support learning
	Assessment guiding instructional planning	Adapting or planning instruction according to assessment
Assessment Practices	Summative assessment dominating teaching practice	Teaching or planning according to the requirements of assessment
	Preference for regular written exams	Preference of traditional paper-pen exams over alternative methods
	Preference for alternative assessment	Preference of alternative assessment over traditional methods
	Teacher autonomy in formative assessment decisions	Being able to adapt or redesign formative assessment tools
	Hesitation toward alternative assessment	Unwillingness to use alternative assessment methods due to lack of confidence or knowledge
	Confidence in using alternative assessments	Feeling capable of implementing an alternative assessment
	Assessment outcomes shaping instruction	Using assessment results to adapt teaching
Assessment fairness	Efforts to ensure fairness in assessment	Adapting, interpreting, or designing assessments in various ways to ensure fairness
	Transparency in assessment criteria	Openly sharing the assessment criteria with the students
Assessment Challenges	Time constraints	Experiencing difficulties due to limited time
	Institutional/policy constraints	Institutional or policy-related factors that hinder assessment practices
	Negative impact on teacher well-being	Assessment practices that contribute to or cause teacher burnout
Need for professional development	Need for hands-on and collaborative training	Showing desire for more practical and interactive training

G. TURKISH SUMMARY

1. GİRİŞ

Ölçme ve değerlendirme sınıf içi uygulamalarında her zaman önemli bir rol oynamış ve öğrenci gelişimini takip etmek, öğretimi desteklemek ve öğrencilere ve diğer paydaşlara geribildirim vermek amaçlarıyla kullanılır (Phye, 1996; Popham, 2013). Etkili ölçme değerlendirme öğretmenlerin güçlü bir ölçme değerlendirme okuryazarlığına sahip olmasını gerektirir. Bu konu son yirmi yılda çok daha fazla ilgi görmeye başlamıştır (DeLuca vd., 2016a).

Ölçme değerlendirme okuryazarlığının tanımı önemli ölçüde değişmiştir (DeLuca vd., 2019). İlk başlarda öğretmenlerin teknik ve profesyonel becerilerine odaklanmış, sonrasında öğretme uygulamalarını şekillendiren inançlar ve düşünelere odaklanmıştır (Opre, 2015). Bu inançlar genellikle öğretmenlerin kişisel değerlerine bağlıdır ve öğretmenlerin ölçme değerlendirmeyi nasıl uyguladıklarını ve yorumladıklarını etkiler (Skott, 2015). Opre (2015) öğretmen inançlarını 4 kategoriye ayırmıştır:

1. Öğretim ve öğrenme için bir rehber olarak ölçme.
2. Öğrencilerin öğrenmelerinin sorumluluğunu üstlenmelerinin bir yolu olarak ölçme.
3. Öğretim ve öğrenmede hesap verebilirlik aracı olarak ölçme.
4. Etkisiz ölçmenin öğretimi, planlamayı ve öğrenenleri olumsuz etkilemesi.

Dil değerlendirme okuryazarlığı öğretmenlerin değerlendirme araçlarını etkili bir şekilde tasarlaması, uygulaması ve sonuçlarını yorumlamasına hizmet eden bir beceriler bütünü olarak tanımlanır (Coombs vd., 2020). Dil değerlendirme okuryazarlığı dil öğretmenlerinin sahip olması gereken temel özelliklerden biridir ve araştırmalarda üzerinde oldukça durulmaktadır. Ancak Türkiye’de, ölçme

okuryazarlığı ile sınıf içi uygulamalar arasında bir boşluk vardır (Öz & Atay, 2017); bu durum, kısmen geleneksel ölçmeden alternatif ölçmeye geçişte yaşanan zorluklardan kaynaklanmaktadır (Tierney, 2006). Birçok bu durumun yaşanmasında eğitim eksikliği, alternatif değerlendirme araçları konusunda yeterince tecrübenin olmaması ve birçok farklı değerlendirme aracını kullanmada yaşanan zorluklar gibi etmenlerin rol oynadığını söyler (DeLuca vd., 2019; Pastore, 2023). Ancak bu sorun aynı zamanda öğretmen eğitiminde ölçmeye sınırlı düzeyde önem verilmesinden ve mesleki gelişim fırsatlarının yetersizliğinden kaynaklanmaktadır (Vogt vd., 2008; Büyükkaracı, 2016). Örneğin, Türkiye’de öğretmen adayları ölçme değerlendirmeyle ilgili bir ders almaktadır (Hatipoğlu, 2015). Tüm bu zorluklara rağmen, öğretmenlerin ölçme ve değerlendirmenin önemi konusundaki farkındalıkları yüksektir ve ölçme ve değerlendirme İngilizce öğretiminin en temel unsurlarından biri olmaya devam etmektedir.

Bu araştırmada DeLuca vd. (2016) tarafından geliştirilen ve ölçme değerlendirme okuryazarlığını 4 boyutta kavramsallaştıran Sınıf-içi Değerlendirmeye Yönelik Yaklaşımlar Envanteri çerçevesi kullanılmaktadır. Bu çerçeveye göre ölçme değerlendirmenin 4 boyutu şöyledir:

1. Ölçme ve değerlendirmenin amacı: Öğrenme amacına hizmet eden ölçme ve değerlendirme araçlarının seçilmesi.
2. Ölçme ve değerlendirmenin süreci: Ölçme ve değerlendirme araçlarının tasarlanması, uygulanması ve sonuçlarının yorumlanması.
3. Ölçme ve değerlendirmede adalet: Ölçme ve değerlendirmede adaletin sağlanması.
4. Ölçme ve değerlendirmenin kuramı: Ölçme anlayışını şekillendiren kavramsal ve pedagojik etkenlerin göz önünde bulundurulması.

Bu çerçeve öğretmenlerin ölçme ve değerlendirmeye olan yaklaşımları konusunda çok yönlü ve geniş bir açıdan baktığı için tercih edilmiştir.

1.2 Çalışmanın amacı ve araştırma soruları

Bu çalışma, Ankara’daki özel bir okulda görev yapan İngilizce öğretmenlerinin ölçme yaklaşımlarını, incelemeyi, öğretmenlik deneyiminin ölçme yönelimleri

üzerindeki etkisini arařtırmayı ve öđretmenlerin sınıf ii ölçme algılarını ve deneyimlerini incelemeyi amaçlamaktadır.

Arařtırma soruları řunlardır:

1. İngilizce öđretmenleri, ACAI çerçevesinde belirlenen dört ölçme teması (ölçmenin amacı, ölçme süreci, ölçmede adalet ve ölçme kuramı) bağlamında belirli bir sınıf ii ölçme yaklaşımını ne ölçüde desteklemektedir?
2. Öđretmenlik deneyimi, İngilizce öđretmenlerinin ACAI çerçevesinde belirlenen dört ölçme teması (ölçmenin amacı, ölçme süreci, ölçmede adalet ve ölçme kuramı) bağlamında belirli bir sınıf ii ölçme yaklaşımını desteklemesini nasıl şekillendirmektedir?
3. İngilizce öđretmenleri sınıf ii ölçmeyi nasıl algılamakta ve buna ilişkin deneyimlerini nasıl yansıtmaktadır?

1.3 Çalışmanın önemi

Ölçme, öđretim ve öğrenmenin temel bir unsurudur ve öđretmenler bu alana önemli ölçüde zaman ayırmaktadır (Stiggins, 1999). Öđretmenler ölçmenin önemini kabul etseler de, pek çok öđretmen ölçme okuryazarlığında kendini yetersiz hissetmekte (Berry ve diđerleri, 2019) ve öğrenci performansını deđerlendirmeye hazır olmadığını düşünmektedir (Mertler, 2003). Öđretmen yetiřtirme programları çođu zaman yeterli ölçme eğitimi sunmamakta, kurumsal baskılar ise etkili uygulamaları daha da zorlařtırmaktadır (Hatipođlu, 2015). Bu zorluklar göz önünde bulundurulduğunda, ölçme okuryazarlığı konusunda farkındalık yaratmak ve sınıf ortamına uygulanabilir, pratik ölçme yöntemleri sunmak önemlidir (Willis vd., 2013; Coombs vd., 2018). Bu çalışma, İngilizce öđretmenlerinin algılarını ve uygulamalarını inceleyerek öđretmenler, okul yöneticileri ve mesleki gelişim planlayıcıları ve öđretmen eğitmenleri için katkı sağlamayı amaçlamaktadır.

Ayrıca, Türkiye’de özel okul İngilizce öđretmenlerinin ölçme yaklaşımlarını ACAI çerçevesi aracılığıyla inceleyen bir çalışma olmaması nedeniyle literatürdeki bir boşluğu da doldurmaktadır. Kiřisel ve kurumsal faktörlerin ve öđretim deneyiminin sınıf ii ölçmeyi nasıl şekillendirdiđini analiz ederek özel eğitim kurumlarında ölçme okuryazarlığının anlaşılmasına katkı sunmaktadır.

3. YÖNTEM

3.1 Araştırma deseni

Bu açıklayıcı durum çalışması, Ankara'daki özel bir okulda görev yapan İngilizce öğretmenlerinin ölçme yaklaşımlarını araştırmak için gömülü karma yöntem yaklaşımını kullanmıştır. Çalışma, öğretim deneyiminin ölçme yönelimlerini ve algılarını nasıl etkilediğini incelemiştir. Durum çalışması tasarımı, ölçme yönelimlerinin gerçek dünya bağlamında nasıl ve neden oluştuğunu araştırmaya olanak tanıdığı için seçilmiştir (Yin, 2018). Veriler ACAI anketinin uyarlanmış hali (DeLuca vd., 2016) ve yarı yapılandırılmış görüşmeler aracılığıyla elde edilmiştir ve nicel ve nitel veriler birleştirilmiştir.

3.2 Araştırma ortamı

Çalışma, Ankara'daki büyük bir özel okulun ortaokul İngilizce zümresinde gerçekleştirilmiştir. Bu kurumun ülke genelinde birden fazla kampüsü ve 7.000'in üzerinde öğrencisi bulunmaktadır. İngilizce zümresi 24 öğretmenden oluşmakta, karma yeterlilik düzeyindeki (A2–B2) sınıflarda haftada 10 saat İngilizce dersi verilmektedir.

Ölçme uygulamaları kurumsal olarak yapılandırılmıştır. Özetleyici değerlendirme araçları ölçme ve değerlendirme birimi tarafından hazırlanıp tüm kampüslerde uygulanmaktadır. Öğretmenler bu araçlara geri bildirim sunabilmektedir ancak genel anlamda özetleyici değerlendirmelerde sınırlı özerkliğe sahiptirler. Yazma sınavlarında örnek değerlendirme kalibrasyonu, kör çift değerlendirme ve rastgele kontrol ile standart puanlama sağlanmaktadır. Biçimlendirici değerlendirmeler öğretmenler tarafından tasarlanırsa da kurumsal hedeflerle uyumlu olmakta ve bölüm başkanı tarafından onaylanmaktadır.

Ölçme ve değerlendirme uygulamaları gibi mesleki gelişim de kurumsal olarak yapılandırılmıştır. Deneyimli öğretmenler yılda en az üç çevrimiçi eğitim, yeni öğretmenler ise altı eğitim almakta; ayrıca gözlemler, akran değerlendirmeleri ve öz

değerlendirmeler yapılmaktadır. Her gözlemin ardından geribildirim ve hedef belirleme süreci takip edilmektedir. Bu yapılandırılmış ölçme ortamı derinlemesine bir durum çalışması için uygun bir bağlam oluşturmuştur.

3.3 Katılımcılar

Uygun 21 öğretmenden 18'i çalışmaya katılmıştır. Amaçlı ölçüt örnekleme ile en az beş yıl öğretmenlik ve bir yıl kurumda çalışma deneyimine sahip olan katılımcılar seçilmiştir. Katılımcılar orta deneyimli (5–10 yıl) ve yüksek deneyimli (11–15 yıl) olarak kategorize edilmiştir. Zümre içerisinde mesleki deneyime sahip olmayan öğretmen olmadığı için bu grup çalışmaya dahil edilememiştir.

3.4. Veri toplama

Nicel veri öğretmenlerin farklı ölçme yaklaşımlarını destekleme düzeylerini değerlendirmek için kullanılan ACAI v3.0 anketi (senaryo temelli, Likert ölçekli) aracılığıyla toplanmıştır. Anketin birinci bölümü araştırmanın amaçlarına göre düzenlenmiş ve üçüncü bölümü ise araştırma amaçlarına hizmet etmediğinden dolayı kullanılmamıştır. Senaryo temelli sorulardan oluşan ikinci bölümü ise DeLuca vd. (2016) tarafından tasarlandığı şekliyle kullanılmıştır. Pilot uygulamanın ardından anketler basılı olarak dağıtılmış ve 18 katılımcının tamamından toplanmıştır.

Nitel veri yarı yapılandırılmış görüşmeler aracılığıyla toplanmıştır. Yarı yapılandırılmış görüşme soruları Çayır'dan (2022)den uyarlanmıştır. Bu görüşmeler yardımıyla araştırmacı öğretmenlerin sınıf içi ölçmeye ilişkin algılarını, karşılaştıkları zorlukları ve ölçme felsefelerini incelemiştir. Sorular İngiliz Dili Öğretimi bağlamına uyarlanmış ve soruların anlaşılabilirliğinden emin olabilmek adına zümre içerisindeki çalışmaya katılmayan bir öğretmen ile pilot görüşme uygulaması yapılmıştır. 17 katılımcıyla Zoom platformu aracılığıyla veya yüz yüze görüşme yapılmıştır. Görüşmeler ortalama 22 dakika sürmüştür ve izin alınarak kaydedilmiştir.

3.5 Veri analizi

Nicel veriler SPSS 23 kullanılarak betimsel istatistiklerle analiz edilmiştir. Nitel veriler ise MAXQDA 2024 aracılığıyla tümevarımsal tematik analiz (Braun & Clarke, 2006) tekniğiyle analiz edilmiştir. Görüşme dökümleri kodlanmış, temalara ayrılmış, rafine edilmiş ve temsil edici alıntılarla desteklenmiştir.

4. BULGULAR

Anket yanıtlarının analizi İngilizce öğretmenlerinin ACAI temaları kapsamında farklı sınıf içi ölçme yaklaşımlarına yönelik destek düzeylerinde çeşitlilik gösterdiğini ortaya koymuştur. Ölçmenin amacı temasında, üç yaklaşım da görece yüksek düzeyde onay almış, ancak en güçlü destek öğrenme için ölçmeye (AfL) yönelmiştir (O= 4.50, SS= 0.81). Bu sonuç, öğretmenlerin ölçmenin öğretimi yönlendirmek ve öğrenme çıktılarının geliştirilmesi amaçlarıyla kullanılması gerektiğini düşündüklerini göstermektedir. Ölçme süreci temasında katılımcılar üç yaklaşımı da güçlü biçimde desteklemiş, bu da ölçmenin dikkatle planlanması, uygulanması ve iletişimine verdikleri değeri yansıtmaktadır. En yüksek destek tasarım yaklaşımına verilmiştir (O= 5.22, SS= 0.77). Bu, öğretmenlerin yüksek nitelikli ölçme araçları geliştirmeye verdikleri önemi göstermektedir. Ölçmede adalet temasında ise daha zıt bir tablo ortaya çıkmıştır; üç yaklaşımın hiçbiri yüksek destek için 4.0 eşiğine ulaşmamıştır. Ölçmenin kuramı temasında ise öğretmenler üç yaklaşımı da yüksek düzeyde desteklemiş, bu da ölçme ilkelerine dair geniş bir kuramsal anlayışları olduğunu göstermektedir. Tutarlı kuram (O= 4.57, SS= 0.84) ve bağlamsal kuram (O= 4.54, SS= 0.59) neredeyse eşit puanlanmıştır. Bu da öğretmenlerin hem ölçme standartlarını tutarlı biçimde uygulamanın hem de bağlama göre esneklik göstermenin önemini takdir ettiklerini ortaya koymaktadır.

Öğretim deneyimine göre analiz edildiğinde, iki grup arasında dikkat çekici farklar ortaya çıkmıştır. Orta düzeyde deneyimli öğretmenler (5–10 yıl), ölçme amacı temasındaki üç yaklaşımın tümünü yüksek düzeyde desteklemiş ve en yüksek puanı öğrenme için değerlendirme yaklaşımına vermiştir (O= 4.61). Benzer şekilde, ölçme süreci yaklaşımlarının da tümünü desteklemişlerdir ve özellikle tasarım (O= 5.11) ve iletişim (O= 4.75) yaklaşımları öne çıkmıştır. Ayrıca, adalet temasında

farklılaştırılmış adaleti yüksek desteklemiş tek grup olmuşlardır (O= 4.14). Kuram temasında ise tüm yaklaşımları yüksek puanlamış, en yüksek değeri tutarlı kurama vermişlerdir. (O= 4.86).

Yüksek deneyimli öğretmenler (11–15 yıl) ise ölçme amacı temasında yalnızca öğrenme için değerlendirme yaklaşımını güçlü biçimde desteklemiştir (O= 4.39). Ölçme sürecinde tasarım (O= 5.33) ve iletişim (O= 4.58) yaklaşımları öne çıkmıştır. Adalet yaklaşımlarından hiçbiri güçlü destek almamıştır. Kuram temasında ise üç yaklaşımı da destekleseler de en yüksek puanı bağlamsal kurama vermişlerdir (O= 4.47). Genel olarak iki grup da öğrenme için değerlendirme, ölçme tasarımı ve açık iletişim yaklaşımlarını önceliklendirmiştir. Ancak orta düzeyde deneyimli öğretmenler, temalara genel olarak daha fazla destek göstermiştir.

Nitel bulgular, İngilizce öğretmenlerinin ölçmeyi çok boyutlu ve çeşitli öğretim amaçlarına hizmet eden bir süreç olarak algıladıklarını göstermiştir. Tüm deneyim seviyelerinde katılımcılar, ölçmeyi öncelikle öğrenci ihtiyaçlarını belirlemek, ilerlemeyi izlemek ve öğretim kararlarını yönlendirmek için kullandıklarını ifade etmişlerdir. Ölçmeyi hem tanılayıcı hem de geliştirici bir araç olarak görmüşler. Ölçme ve değerlendirmenin kendilerine öğrenme açıklarını belirleme, hedefe yönelik geri bildirim sağlama ve ders planlarını buna göre uyarlama imkânı sunduğunu vurgulamışlardır. Ayrıca, öğrencilerin kendi ilerlemelerini değerlendirmeleri, hedef belirlemeleri ve öğrenme süreçlerine daha aktif katılmaları için ölçmeyi bir öğrenme aracı olarak da görmüşlerdir. Ölçme sonuçlarının günlük ders planlarının yanı sıra uzun vadeli müfredat düzenlemelerini de etkilediği ve öğretmenlerin gelecek akademik yıllara ilişkin kararlarını bu verilere dayandırdıkları belirtilmiştir.

Ölçme uygulamaları bağlamında, katılımcılar öğretim ortamının ağırlıklı olarak özetleyici değerlendirmeler tarafından şekillendirildiğini ifade etmişlerdir. Ders içeriklerinin ve hızının genellikle sınav gereklilikleriyle uyumlu olduğunu belirtmişlerdir. Bazı orta deneyimli öğretmenler, güvenilirlik ve aşinalık gereklilikleriyle geleneksel yazılı sınavlara öncelik verdiklerini söylediler de, çoğu katılımcı akran ve öz değerlendirme, portfolyo, projeler ve sınıf içi gözlemler gibi alternatif ölçme yöntemlerini tercih ettiklerini belirtmiştir. Bu alternatifler, daha derin

öğrenmeyi, iş birliğini ve üstbilişsel farkındalığı desteklediği için değerli bulunmuştur. Ancak bazı öğretmenler hem geleneksel hem de alternatif araçları bir arada kullanan karma bir yaklaşımı tercih etmiştir. Belirgin bir ayrışma noktası ise alternatif ölçme yöntemlerinin uygulanmasındaki güven olmuştur. Çoğu öğretmen çeşitli biçimlendirici araçları kullanmada kendini yetkin hissederken, özellikle bazı orta deneyimli öğretmenler sınırlı eğitim, deneyim eksikliği veya proje çalışmalarında akademik dürüstlük gibi sorunlar nedeniyle güven eksikliği yaşadıklarını belirtmiştir. Buna rağmen, öğretmenler biçimlendirici ölçmeleri tasarlama ve uyarlama konusunda yüksek derecede özerkliğe sahip olduklarını, öğrenenlerin ihtiyaçlarına göre anlık değişiklikler yaptıklarını ve farklı yeterlilikleri karşılamak için araçları farklılaştırdıklarını ifade etmişlerdir. Adalet konusu öğretmenler için temel bir kaygı olarak ortaya çıkmıştır. Katılımcılar, adaleti sağlamak için çeşitli stratejiler uyguladıklarını belirtmişlerdir. Bireysel öğrenme profillerine göre değerlendirme araçlarını uyarlama, görev ve araçları farklılaştırma, tutarlılık için meslektaşlarla iş birliği yapma ve standardizasyon ile çift değerlendirmenin uygulanması bu stratejiler arasında yer almaktadır. Ayrıca şeffaflık da önemli görülmüş; öğretmenler rubriklerin ve ölçütlerin öğrencilere açıkça paylaşılmasının beklentileri netleştirdiğini ve adaleti artırdığını vurgulamışlardır.

Zorluklar arasında en sık dile getirilen engel zaman kısıtlamaları olmuştur. Bu durum, öğretmenlerin çeşitli değerlendirmeleri tasarlama, zamanında geri bildirim sağlama ve sonuçlara göre harekete geçme kapasitelerini sınırlamaktadır. Kurumsal kısıtlamalar özellikle ölçmelerin zümrenin ölçme ve değerlendirme birimi tarafından merkezi şekilde hazırlanması öğretmen özerkliğini azaltan bir unsur olarak görülmüştür. Orta deneyimli öğretmenler, ölçme ve değerlendirmenin esenlikleri üzerinde olumsuz etki yarattığını, yorgunluk ve zaman baskısı duygusuyla karşı karşıya kaldıklarını daha dile getirmişlerdir.

Mesleki gelişim ihtiyacı hemen tüm katılımcılar tarafından güçlü bir şekilde ifade edilmiştir. Öğretmenler, doğrudan bağlamlarına uygulanabilir, işbirlikçi ve uygulamalı eğitimler talep etmişlerdir. Tek seferlik atölyeler yerine uzun vadeli ve sürekli programların önemini vurgulamışlardır. Bu tür programların kuramsal bilgi

ile sınıf içi gerçeklik arasındaki boşluğu kapatacağı ve öğretmenlerin farklı ölçme uygulamalarında güven ve yetkinlik kazanmalarını sağlayacağı belirtilmiştir.

5. TARTIŞMA VE SONUÇ

Bu çalışma, özel bir okulda görev yapan İngilizce öğretmenlerinin ölçme-değerlendirme yaklaşımlarını, öğretim deneyiminin bu yaklaşımlara etkisini ve sınıf içi ölçme-değerlendirme anlayışlarını incelemiştir. Bulgular, ACAI çerçevesinin dört teması üzerinden analiz edilmiştir.

Ölçme-değerlendirmenin amacı açısından tüm öğretmenler öğrenme için değerlendirme yaklaşımını benimsemiş ve böylece biçimlendirici değerlendirmenin öğretim ve öğrenmeye katkısının farkında olduklarını göstermiştir. Ayrıca, değerlendirmeyi öğrenci ihtiyaçlarını belirlemek, ilerlemeyi izlemek, öğretimi planlamak ve öğrencilerin kendi performanslarını değerlendirip hedef koymalarını sağlamak için kullanmışlardır. Bu durum öğrenme olarak değerlendirme yaklaşımına işaret etmektedir. Bununla birlikte, kurumsal zorunluluklar nedeniyle öğrenmenin değerlendirilmesi uygulamaları da gerekli görülmüştür. Genel olarak öğretmenler, araçlarını bağlama ve öğrenci ihtiyaçlarına göre uyarladıkları ve değerlendirmeyi mesleki iyi oluşlarıyla ilişkilendirdikleri için sosyokültürel bir ölçme anlayışına sahiptirler.

Ölçme-değerlendirme süreci bağlamında öğretmenler en çok geçerli ve güvenilir araçlar tasarlamaya önem vermişlerdir. Ayrıca öğrenciler ve velilerle açık iletişimi vurgulamışlardır. Hemen hemen tüm öğretmenlerin öğrenci ihtiyaçlarına göre biçimlendirici araçları uyarladığı ortaya konulmuş ve bu da kurumsal kısıtlamaların özellikle özetleyici değerlendirmelerde özerkliği sınırladığını göstermiştir.

Ölçme-değerlendirmede adalet konusunda öğretmenlerin yanıtları önceki araştırmalarla tam olarak örtüşmemiştir. Yine de adaleti sağlamak için öğrenci gelişimini dikkate alma, farklılaştırılmış araçlar kullanma, öğretmen kanaatiyle notları destekleme, öğretmenler arası iş birliğiyle güvenilirliği artırma ve şeffaf rubrikler paylaşma gibi uygulamalar öne çıkmıştır. Ancak öğretmenlerin adalet

anlayışları çoğunlukla tutarlılık ve eşitlik kavramlarıyla sınırlı kalmıştır. Bu durum olası bir terminoloji farkındalığı eksikliğine işaret etmektedir.

Ölçme-değerlendirme kuramı açısından öğretmenler en çok tutarlı ve bağlamsal kuramı desteklemişlerdir. Yani güvenilirlik ve bağlama uygunluğu vurgulamışlardır. Değerlendirmeyi ek bir unsurdan çok öğretimle bütünleşmiş bir süreç olarak görmüşlerdir. Ancak dengeli kuram görece düşük destek almıştır. Bu da geçerlik, güvenilirlik ve adaleti dengeleme konusunda yeterli eğitim eksikliğine bağlanabilir.

Her ne kadar öğretmenler öz-değerlendirme, akran değerlendirmesi, proje ve portfolyo gibi alternatif yöntemleri tercih etseler de kurumsal baskılar nedeniyle sınav odaklı uygulamalar hala baskındır. Bazı öğretmenler özellikle kolay uygulanabilir olduğu için geleneksel sınavları tercih ederken, diğerleri alternatif yöntemleri öğretim için vazgeçilmez görmüştür. Bununla birlikte, alternatif yöntemlere dair sınırlı eğitim, yapay zekâ kullanımına ilişkin kaygılar ve kurumsal düzenleme eksiklikleri dile getirilmiştir.

Öğretmenlerin belirttiği zorluklar arasında en sık yinelenenler değerlendirmeyi planlama ve geri bildirim için zaman yetersizliği, adaleti sınırlayan kurumsal kısıtlamalar ve iş yüküne bağlı stres olmuştur. Daha az deneyimli öğretmenler özellikle ölçme ve değerlendirmenin duygusal yükünü daha çok dile getirmiştir. Bu da öğretmen yetiştirme programlarında son dönemde mesleki esenliğe daha fazla vurgu yapılmasıyla açıklanabilir.

Son olarak, öğretmenler işbirlikçi, uygulamaya dönük ve sürdürülebilir mesleki gelişim fırsatlarına güçlü bir ihtiyaç duyduklarını belirtmişlerdir. Bu durum sınav odaklılıktan uzaklaşıp daha bütüncül ve biçimlendirici yaklaşımları destekleyen kurumsal reformlara duyulan acil gereksinime işaret etmektedir.

H. THESIS PERMISSION FORM / TEZ İZİN FORMU

(Please fill out this form on computer. Double click on the boxes to fill them)

ENSTİTÜ / INSTITUTE

- Fen Bilimleri Enstitüsü** / Graduate School of Natural and Applied Sciences
- Sosyal Bilimler Enstitüsü** / Graduate School of Social Sciences
- Uygulamalı Matematik Enstitüsü** / Graduate School of Applied Mathematics
- Enformatik Enstitüsü** / Graduate School of Informatics
- Deniz Bilimleri Enstitüsü** / Graduate School of Marine Sciences

YAZARIN / AUTHOR

Soyadı / Surname : Çetin

Adı / Name : Gözdenur

Bölümü / Department : İngiliz Dili Öğretimi / English Language Teaching

TEZİN ADI / TITLE OF THE THESIS (İngilizce / English): INVESTIGATING THE ASSESSMENT APPROACHES OF EFL TEACHERS IN A PRIVATE SCHOOL IN ANKARA: IMPACT OF TEACHING EXPERIENCE

TEZİN TÜRÜ / DEGREE: **Yüksek Lisans** / Master **Doktora** / PhD

1. **Tezin tamamı dünya çapında erişime açılacaktır.** / Release the entire work immediately for access worldwide.
2. **Tez iki yıl süreyle erişime kapalı olacaktır.** / Secure the entire work for patent and/or proprietary purposes for a period of **two years**. *
3. **Tez altı ay süreyle erişime kapalı olacaktır.** / Secure the entire work for period of **six months**. *

* Enstitü Yönetim Kurulu kararının basılı kopyası tezle birlikte kütüphaneye teslim edilecektir. /

A copy of the decision of the Institute Administrative Committee will be delivered to the library together with the printed thesis.

Yazarın imzası / Signature

Tarih / Date

(Kütüphaneye teslim ettiğiniz tarih. Elle doldurulacaktır.)

(Library submission date. Please fill out by hand.)

Tezin son sayfasıdır. / This is the last page of the thesis/dissertation.