

**REPUBLIC OF TURKEY**  
**ÇAĞ UNIVERSITY**  
**INSTITUTE OF SOCIAL SCIENCES**  
**DEPARTMENT OF ENGLISH LANGUAGE EDUCATION**

**TEACHER EDUCATORS' CONCEPTIONS ON LANGUAGE ASSESSMENT**  
**LITERACY: A CASE STUDY FROM TÜRKİYE**

**THESIS BY**

**Gamze KAPUCU**

**Supervisor: Prof. Dr. Şehnaz ŞAHİNKARAKAŞ**

**Member of Examining Committee: Dr. Semiha GÜRSOY**

**Member of Examining Committee: Dr. Deniz ELÇİN (Siirt University)**

**MASTER THESIS**

**MERSİN / SEPTEMBER 2023**

**APPROVAL****REPUBLIC OF TURKEY****ÇAĞ UNIVERSITY****DIRECTORSHIP OF THE INSTITUTE OF SOCIAL SCIENCES**

We **certify** that thesis under the title of “Teacher Educators’ Conceptions on Language Assessment Literacy: A Case Study from Türkiye” which was prepared by our student Gamze KAPUCU with number 2020008011 is satisfactory **consensus/by majority of votes** for the award of the degree of **Master of Arts** in the Department of **English Language Education**.

(The original copy is wet signed in the institute directorate)

Univ. Inside permanent member-Supervisor-Head of Examining Committee:  
Prof. Dr. Şehnaz ŞAHİNKARAKAŞ

(The original copy is wet signed in the institute directorate)

Univ. Inside - permanent member:  
Dr. Semiha GÜRSOY

(The original copy is wet signed in the institute directorate)

Univ. Outside - permanent member: Dr. Deniz ELÇİN (Siirt University)

**I confirm that the signatures above belong to the academics mentioned.**

(The original copy is wet signed in the institute directorate)

20 / 09 / 2023

Prof. Dr. Murat KOÇ

Director of Institute of Social Sciences

**Note: The uncited usage of the reports, charts, figures and photographs in this thesis, whether original or quoted for mother sources is subject to the Law of Works of Arts and Thought. No: 5846.**

**ETHICS DECLARATIONS**

<b>Student's</b>	Name & Surname:	Gamze KAPUCU
	Number:	2020008011
	Department:	English Language Education
	Program:	Master Thesis (x) Ph.D. Thesis ( )
	Thesis Title:	Teacher Educators' Conceptions on Language Assessment Literacy: A Case Study from Türkiye

I hereby declare that;

I prepared this master thesis in accordance with Çağ University Institute of Social Sciences Thesis Writing Directive,

I prepared this thesis within the framework of academic and ethics rules,

I presented all information, documents, evaluations and findings in accordance with scientific ethical and moral principles,

I cited all sources to which I made reference in my thesis,

The work of art in this thesis is original,

I hereby acknowledge all possible loss of rights in case of a contrary circumstance. (in case of any circumstance contradicting with my declaration)

Gamze KAPUCU

## DEDICATION

*To my family...*

## ACKNOWLEDGEMENTS

To my supervisor, Prof. Dr. Şehnaz ŞAHİNKARAKAŞ, I would like to express my sincerest and utmost gratitude for the continuous support, patience, understanding, and encouragement she has shown me from the very beginning to the end of my master's education. I am deeply indebted to her tireless efforts at every stage of my thesis. This thesis would not have seen the light of day without her.

I extend my thanks to the members of my thesis committee, Dr. Deniz ELÇİN and Dr. Semiha GÜRSOY, for their valuable feedback and contributions to improving my thesis. My appreciation goes to the faculty members of Çağ University, particularly Prof. Dr. Jülide İNÖZÜ, Dr. Senem ZAIMOĞLU, Dr. Aysun DAĞTAŞ, and Dr. Betül ALTAŞ, for generously sharing their knowledge with us.

I thank the valuable faculty members who have shown patience while I was collecting the data I needed for this research.

I am grateful to the faculty members of Afyon Kocatepe University, Assoc. Prof. Dr. Fatih GÜNGÖR, Assoc. Prof. Dr. Cahit ERDEM and Dr. Onur ULUDAĞ for their invaluable support during this process.

To my precious friends, Şule EMİRZEOĞLU, İklima ECE, and Gülenay AKTEPE, thank you for never losing faith in me. Your belief in me has kept my spirits and motivation high during this process. Words cannot express my gratitude to Merve ARABACI for her generous support, late-night sessions, and encouragement. I am also grateful to Barış VAROL for his unwavering support and valuable scientific contributions throughout this challenging journey.

Finally, I want to express my heartfelt thanks to my beloved family for their unconditional love, sacrifices, and continuous support.

**ABSTRACT****TEACHER EDUCATORS' CONCEPTIONS ON LANGUAGE ASSESSMENT  
LITERACY: A CASE STUDY FROM TÜRKİYE****Gamze KAPUCU****Master Thesis, Department of English Language Education****September 2023, 104 Sayfa****Supervisor: Prof. Dr. Şehnaz ŞAHİNKARAKAŞ**

This study investigated how teacher educators in Türkiye conceptualize Language Assessment Literacy and the relevancy of Taylor's (2013) Language Assessment Literacy model in the Turkish context. This qualitative research was carried out with faculty members working in English Language Teaching departments of various universities in Türkiye. Data collection was conducted using a semi-structured interview guide and closed-response questionnaire items developed by Bøhn and Tsagari (2021). According to the findings of this study, the perceptions of teacher educators regarding Taylor's Language Literacy Model have revealed a framework comprising theoretical knowledge of assessment, competence in scoring and statistics, language pedagogy knowledge, content knowledge, and social context knowledge. The research findings suggested the need for establishing a comprehensive Language Assessment Literacy culture within the language assessment field.

**Keywords:** Language assessment, language assessment literacy, foreign language teaching, teacher education

**ÖZ****ÖĞRETMEN EĞİTİCİLERİNİN DİL DEĞERLENDİRME OKURYAZARLIĞINA İLİŞKİN ANLAYIŞLARI: TÜRKİYE'DEN BİR DURUM ÇALIŞMASI****Gamze KAPUCU****Yüksek Lisans Tezi, İngiliz Dili Eğitimi Anabilim Dalı****Tez Danışmanı: Prof. Dr. Şehnaz ŞAHİNKARAKAŞ****Eylül 2023, 104 Pages**

Bu çalışmada, Türkiye'deki öğretmen eğitimcilerinin Dil Değerlendirme Okuryazarlığı kavramını nasıl anlamlandırdıkları incelenmiş ve Taylor'ın (2013) Dil Değerlendirme Okuryazarlık modelinin Türkiye bağlamında değerlendirilmesi yapılmıştır. Bu nitel çalışma, Türkiye'deki çeşitli üniversitelerin İngiliz Dili Eğitimi bölümlerinde görev yapan öğretim elemanlarıyla gerçekleştirilmiştir. Veri toplama aracı olarak Bøhn ve Tsagari (2021) tarafından hazırlanan yarı yapılandırılmış mülakat soruları ve anket formu kullanılmıştır. Çalışmanın bulguları değerlendirildiğinde, Taylor'ın (2013) Dil Okuryazarlık Modeli bağlamında Türkiye'deki öğretmen eğitimcilerinin görüşleri; değerlendirmeye yönelik teorik bilgi, puanlama ve istatistik bilgisi, pedagojik dil bilgisi, alan bilgisi ve sosyal bağlam bilgisi alt boyutlarından oluşan bir model ortaya koymuştur. Araştırma sonucunda, dil eğitimi alanında kapsamlı bir Dil Değerlendirme Okuryazarlığı kültürünün oluşturulmasına duyulan gereksinime yönelik öneriler sunulmuştur.

**Anahtar Kelimeler:** Dil değerlendirme, dil değerlendirme okuryazarlığı, yabancı dil öğretimi, öğretmen eğitimi

## TABLE OF CONTENTS

COVER.....	i
APPROVAL .....	ii
ETHICS DECLARATIONS.....	iii
DEDICATION.....	iv
ACKNOWLEDGEMENTS.....	v
ABSTRACT.....	vi
ÖZ .....	vii
TABLE OF CONTENTS.....	viii
ABBREVIATIONS .....	x
LIST OF TABLES.....	xi
LIST OF FIGURES .....	xii
LIST OF APPENDICES.....	xiii
1. INTRODUCTION .....	1
1.1. Statement of the Problem .....	1
1.2. Significance of the Study.....	2
1.3. Purpose of the Study .....	3
1.4. Conceptual Framework and Related Studies .....	4
1.4.1. Conceptualizing Language Assessment Literacy.....	4
1.4.2. Language Assessment Literacy Models .....	7
1.4.3. Taylor’s (2013) Stakeholder Model for Classroom Teachers .....	15
1.4.4. Research on Classroom Teachers’ LAL.....	20
1.4.5. Reports from the Teaching and Learning International Survey .....	25
2. METHODOLOGY .....	30
2.1. Research Design.....	30
2.2. Participants .....	30
2.3. Data Collection Tools.....	32
2.4. Data Collection.....	35
2.5. Data Analysis .....	35



3. FINDINGS.....	39
3.1. Teacher Educators’ Conceptions on LAL.....	39
3.2. The Relevancy of Taylor’s (2013) LAL Model for Turkish Teacher Educators.....	55
4. DISCUSSION AND CONCLUSION .....	66
4.1. Classroom Teachers’ LAL Profile from Teacher Educators’ Perspectives.....	66
4.2. Comparing Taylor’s (2013) and Turkish Educators’ Perspectives on Classroom Teachers’ LAL Profiles.....	72
REFERENCES .....	75
APPENDICES .....	82

## ABBREVIATIONS

- EFL** : English as a Foreign Language  
**ELT** : English Language Teaching  
**LAL** : Language Assessment literacy  
**MoNE** : Ministry of National Education  
**CEFR** : Common European Framework of Reference for Languages

## LIST OF TABLES

<b>Table 1.</b> Teachers' Classroom Assessment Practices from TALIS 2018 Reports .....	29
<b>Table 2.</b> Teachers' Classroom Assessment Practices from TALIS 2018 Reports .....	32
<b>Table 3.</b> Operationalization of Taylor's LAL Dimensions .....	34
<b>Table 4.</b> Duration and Mode of the Interviews.....	35
<b>Table 5.</b> Overview of Codes and Themes Incorporating Taylor's (2013) LAL Dimensions .....	37
<b>Table 6.</b> Six-point Likert scale Scoring System .....	55
<b>Table 7.</b> The Highest Ranked Items from Closed-ended Questions .....	56
<b>Table 8.</b> The Lowest Ranked Items from Closed-ended Questions .....	57
<b>Table 9.</b> Knowledge of Theory Dimension .....	58
<b>Table 10.</b> Technical Skills Dimension.....	58
<b>Table 11.</b> Principles and Concepts Dimension .....	59
<b>Table 12.</b> Language Pedagogy Dimension .....	60
<b>Table 13.</b> Sociocultural Values Dimension.....	61
<b>Table 14.</b> Local Practices Dimension .....	62
<b>Table 15.</b> Personal Beliefs and Attitudes Dimension .....	63
<b>Table 16.</b> Scores and Decision-Making Dimension .....	63
<b>Table 17.</b> Taylor's (2013) LAL Levels for Classroom Teachers .....	64
<b>Table 18.</b> Conversion of Average Median Scores .....	64

## LIST OF FIGURES

<b>Figure 1.</b> The Relationship between Construct Validity, Authenticity, and Interactiveness.....	10
<b>Figure 2.</b> Pill and Harding’s (2013) LAL Stages and Descriptions .....	14
<b>Figure 3.</b> Taylor’s (2013) LAL Profile for Classroom Teachers .....	17
<b>Figure 4.</b> Classroom Teacher’s LAL Profiles According to Norwegian Teacher Educators	19
<b>Figure 5.</b> Revised Version of Taylor’s (2013) LAL Profile based on Bøhn and Tsagari (2021) .....	19
<b>Figure 6.</b> Teachers’ Assessment Practices in Participating OECD Countries .....	26
<b>Figure 7.</b> Change in Teachers’ Classroom Assessment Practices from 2013 to 2018 .....	27
<b>Figure 8.</b> Teachers Reporting a High Level of Need for Professional Development in Türkiye .....	28
<b>Figure 9.</b> Classroom Teachers’ LAL Profile from Turkish Teacher Educators’ Perspective .....	65
<b>Figure 10.</b> Comparing Taylor’s (2013) and Turkish Educators’ Perspectives on Classroom Teachers’ LAL Profiles.....	72

**LIST OF APPENDICES**

<b>Appendix A.</b> Approval from the Members of the Ethics Committee .....	82
<b>Appendix B.</b> Consent Form .....	84
<b>Appendix C.</b> Semi-Structured Interview Guide for Teacher Educators .....	85
<b>Appendix D.</b> Closed-ended Questionnaire Form.....	87
<b>Appendix E.</b> Scientific Research and Publication Ethics Board Decision .....	89
<b>Appendix F.</b> Ethical Permission Letter.....	90

## 1. INTRODUCTION

This chapter provides an overview of the problem, purpose, and significance of the study, as well as the underlying assumptions and limitations.

### 1.1. Statement of the Problem

In recent years, assessment in educational settings has been greatly affected by shifting perspectives in language instruction (Sevimel-Şahin & Subaşı, 2021). The transition from traditional assessment approaches and strategies, such as summative and result-oriented assessment, to formative assessment that promotes monitoring learners' development and improving teaching and learning process has become desirable. This shift in preference emerged after the advancements regarding language teaching and learning approaches, which prioritize a learner-centered perspective (Tsagari & Vogt, 2017).

In order to carry out their roles as assessment stakeholders, classroom teachers need to acquire the necessary knowledge and skills to become *assessment literate*. By becoming assessment literate, practitioners will be well-prepared to navigate the complexities of the assessment procedures, make sound judgments about scores and decisions, and promote learning objectives. As Stiggins (2002) advocates, teachers with strong assessment backgrounds cannot be deterred by the challenging technical side of assessment and can incorporate assessment into learning (McMillan & Nash, 2000). Practicing effective classroom assessment enables teachers to make more informed decisions about learners' progress and convey assessment results to students and parents (Brookhart, 1999).

Since teaching and assessment were formerly fragmented into two separate domains, initial teacher education programs neglected language assessment training in favor of methodological instruction (Viengsang, 2016). However, the importance of classroom assessment and teachers' role as assessors were emphasized due to the widespread acceptance of the connection between teaching and assessment. This development has also led to a greater emphasis on teacher education programs to equip teachers with the desired LAL levels (Scarino, 2013; Wach, 2012). In Türkiye, a single Language Testing and Assessment (LTA) course is offered in the last semester of the undergraduate English Language Teaching (ELT) program. Both pre-service and in-service

English as a Foreign Language (EFL) teachers have previously expressed dissatisfaction with the effectiveness of this course, as it fails to provide teacher candidates with the necessary knowledge and expertise for their assessment-related duties (Mede & Atay, 2017). Consequently, teachers often find themselves having to learn about assessment practices on the job (Hatipoğlu, 2015; Vogt & Tsagari, 2014). This situation is further highlighted in the struggle teacher educators face while choosing which aspects of LAL to cover and prioritize in these courses due to time constraints (Inbar-Lourie, 2008).

## **1.2. Significance of the Study**

Innovations in the field of second language and foreign language education have profoundly influenced the approaches employed in classroom-based language testing and assessment. As a result of this development, classroom teachers have assumed a prominent role in the educational assessment domain. Consequently, language teachers' assessment competence and the quality of training provided to them have emerged as a subject of interest. Currently, assessment training courses in pre-service teacher education programs are the primary means by which teachers receive initial LTA training. These courses are intended to equip teachers with the necessary skills to accurately assess language proficiency levels, design effective assessment instruments, provide effective feedback to learners, and make informed decisions about their instruction. As the delivery of these courses relies on teacher educators, they can play a pivotal role in shaping the quality and content of the LTA education provided to prospective EFL teachers.

By exploring teacher educators' understanding of LAL and the applicability of existing LAL frameworks to EFL teachers in Türkiye, this research attempts to add to the expanding body of knowledge about conceptualizing and empirically validating these frameworks. As indicated by Scarino (2013), the development of teacher assessment literacy calls for an inquiry into ways to foster it, one that will engage teachers in a critical analysis of both their own and others' perceptions of the assessment-related issues. Moreover, the way in which different stakeholder groups conceptualize LAL can provide empirical data for policymakers to consider while planning pre-service and in-service training programs (Giraldo, 2020). To the best of this researcher's knowledge, teacher educators' conceptions of LAL have yet to be studied in the context of Türkiye. Therefore, this investigation is

anticipated to yield results that have not been discussed extensively. The findings of this study may also offer insights for EFL instructors and researchers on what aspects of assessment knowledge should be emphasized to improve the language learning process.

### **1.3. Purpose of the Study**

Language assessment is a crucial component of the language teaching and learning process. The importance of LAL has been emphasized globally, given the widespread use of language tests to assess learners' language proficiency levels. The data from these tests are often used to make important decisions, such as student placement, selection and admission, curriculum evaluation and policy making, and so on. At the classroom level, the attention given to these results shifts towards monitoring students' academic progress and planning for the future of their learning trajectories. By developing their LAL competencies, classroom teachers can not only improve their ability to interpret assessment results accurately and provide constructive feedback to students but also become able to make informed decisions about the effectiveness of their teaching. Therefore, developing classroom teachers' LAL skills to ensure that they can design, administer, and interpret language assessments effectively helps promote language learning and teaching processes.

Experienced teachers, adept at evaluating language proficiency and diagnosing learner challenges, provide valuable insights for LAL development, serving as a foundation for deriving universal assessment principles (Brindley, 2001). In the same vein, understanding how stakeholders in education, particularly teacher educators, conceptualize LAL can be helpful for determining the effectiveness of professional development and teacher education training initiatives. Previous studies have highlighted the need for further research incorporating different stakeholders' perspectives in order to establish a comprehensive LAL culture that fits the needs of the language assessment community (Bøhn & Tsagari, 2021; Davies, 2008; Inbar-Lourie, 2008; Sevimeş-Sahin & Subaşı, 2019; Taylor, 2009). Thus, the present study aims to address this gap by exploring how teacher educators perceive classroom teachers' LAL and how their perspectives align with Taylor's (2013) LAL profile for classroom teachers. The following research questions are established for this study to achieve its objectives:

1. How do Turkish teacher educators conceptualize Language Assessment Literacy?



2. What relevance does Taylor's (2013) Language Assessment Literacy model have for Turkish teacher educators?

#### **1.4. Conceptual Framework and Related Studies**

The aim of this section is to look into several interpretations of LAL, the existing LAL models proposed by language assessment researchers, and their implications for educational settings.

##### **1.4.1. Conceptualizing Language Assessment Literacy**

Assessment in education refers to collecting information about learning processes and outcomes, aiming to diagnose learners' current knowledge, evaluate achievement, and improve the quality of teaching and learning (Brown, 2004). Within the field of language education, the purpose of assessment is to determine learners' language ability in a given situation (Bachman & Palmer, 1996, p. 29). Brindley (2001, p. 127) argues that teachers perceive assessment as "an activity which is integrated into the curriculum with the aim of improving learning, rather than a 'one-off' summative event." Therefore, an effective teacher continuously assesses students using many different techniques, including diagnostic assessments, impromptu assessments, self-assessment, observation, and corrective feedback (Brindley, 2001; Brown, 2004). Brown (2004) classifies assessment as "formal" or "informal," depending on how it is used in the classroom. Informal assessment involves unplanned and spontaneous feedback given to students within a classroom environment. This feedback is integrated into activities aimed at eliciting learners' performance. However, the outcomes of these assessments are not utilized to form rigid judgments about learners' competence, nor are they documented. Formal assessments, on the other hand, are specifically designed procedures to assess a wide range of knowledge and skills. They are systematic and deliberate techniques intended to provide teachers and students with a measurement of achievement (Brown, 2004, p. 6). In recent years, there has been an adoption of assessment jargon in language evaluation research discourse (Inbar-Lourie, 2008a, p. 385). Assessment is now considered an inclusive term encompassing "all testing methods and approaches in research or education" (Kunnan, 2004, p. 1). This extensive view dictates the necessity for several ways of assessment to gather data for different objectives in different situations (Huerta-Macias, 1995). The assessment of language knowledge has been acknowledged as a socially constructed activity that is

integrated within the local context. Thus, the involvement of teachers, students, and members of the community has become significant in ensuring assessment partnerships. Moreover, teachers' experiences, assessment beliefs, expectations from their professional community, and external demands have an impact on their assessment practices (Breen et al., 1997; Brindley, 2001; McCallum et al., 1995).

In the current era, stakeholders in the field of education have expediently acknowledged assessment as an important procedure for improving learning. Although the ongoing debate around the paradigms, theories, and approaches that should inform teachers' assessment practices is yet to reach its conclusion, stakeholders involved in education are becoming more aware that there is no one-size-fits-all approach or a universal method for student assessment. This awareness has given rise to the importance of Assessment Literacy (AL) due to the fact that effective assessment implementation relies on various factors at the classroom, school, and system levels (Siarova et al., 2017).

A revision of the competencies required for carrying out assessment in an instructional setting is expected on account of the shifting views surrounding assessment. McMillan (2000) highlights eight principles for assessment that have since become indispensable to teachers and administrators in educational institutions. The aim of this compilation is to provide an extensive repository of knowledge that integrates a variety of perspectives and provides those in the education sector with an overview of the current concerns. The aforementioned principles consist of various disputes that may arise during the course of assessment and affect the way decisions are made. Such disputes may present themselves between formative and summative assessment, criterion- and norm-referenced approaches, traditional and alternative assessment formats, and external standardized testing against classroom assessments. Other elements include formative assessment in education, the value of employing various assessment methods, and the call for equitable, moral, valid, practical, and reliable assessments (McMillan, 2000). To effectively incorporate these principles into practice, it is necessary to become literate in the concepts, skills, and strategies that underlie assessment (Inbar-Lourie, 2008a).

Assessment literacy refers to the capability to comprehend, scrutinize, and put into practice data from learners' performance to promote instructional improvement (Falsgarf,

2005). To gain literacy in assessment means acquiring a practical and theoretical skillset for developing a range of assessment activities, including knowledge of when, why, and how to implement them (Inbar-Lourie, 2008a). Fulcher (2012) defines AL as the knowledge and skills required to select, create, and apply assessments for various purposes. Therefore, AL can equip teachers with the ability to make informed decisions and interpretations about classroom assessments, facilitate assessment-driven learning opportunities for their students, and uphold an effective teaching practice (Stabler-Havener, 2018). According to Inbar-Lourie (2008a), AL is developed in the same manner as initiatives for professional development in socio-cultural pedagogy. The approach taken in both situations is constructivist in nature. Assessment specialists and those involved construct a network of knowledge by debating, analyzing, and probing key subjects pertinent to their environment. The issues presented here are generally considered to be foundational assessment competencies. However, a major point for discussion at this stage is the specific knowledge needed by those involved in conducting language assessments and the degree to which it is included in language assessment courses. Therefore, the fact that these courses are in existence at all suggests that there is a need for other specific competencies in the field of language assessment.

Language Assessment Literacy (LAL) is an interpretation of AL that incorporates language-related elements of assessment, taking into account the complex nature of language instruction, such as linguistic, communicative, and cultural competence (Harding & Kremmel, 2016). Several conceptualizations of LAL have been put forth to identify the competencies assessment stakeholders need to possess (Bachman & Palmer, 1996; Brindley, 2001; Davies, 2008; Fulcher, 2012; Inbar-Lourie, 2008; Pill & Harding, 2013; Taylor, 2013). Bachman and Palmer (1996) define the components of language assessment competence as involving five requirements. First, an awareness of the key variables needs to be taken into account before engaging in language testing activities. This may entail either creating new language tests or choosing one that already exists. The second component involves the knowledge of the core issues and considerations in using language tests appropriately. The third component is related to the knowledge of measurement and evaluation approaches and methods, and the fourth component pertains to the skills required to plan, construct, evaluate, and implement language tests to meet a particular

objective, situation, and audience. Lastly, developing the skills to critically analyze language testing publications is necessary for making informed decisions (Bachman & Palmer, 1996, p. 9).

Pill and Harding (2013) interpreted LAL as the abilities and knowledge required to understand, adequately evaluate, and apply language assessments. Similarly, Tsagari and Vogt (2014, p. 377) defined it as “the ability to design, develop and critically evaluate tests and other assessment procedures, as well as the ability to monitor, evaluate, grade and score assessments on the basis of theoretical knowledge.” Davies (2008) identified two primary LAL branches, namely, “skills” and “knowledge.” The former relates to the ability to create and evaluate assessments, while the latter refers to linguistic expertise and familiarity with assessment techniques. In a different perspective, Inbar-Lourie (2008) found that Davies’ conceptualization neglected the sociocultural aspects of assessment and expanded on this definition by encompassing diverse language education settings and contexts within the constructivist framework of assessment. She emphasized the importance of internalizing local situations and beliefs about assessment and recognizing their influence on learning. Inbar-Lourie (2008a) suggested that any inquiry into LAL should take into account the recent developments in assessment, specifically the increasing adoption of assessment for learning approaches. Teachers, in particular, have a responsibility to carry out classroom assessment activities, keep track of students’ progress in a manner consistent with standards set by external norms, and prepare them for exams administered by external organizations. A similar argument was made by Lam (2014, p. 4), who described LAL as “teachers’ understanding and mastery of assessment concepts, measurement knowledge, test construction skills, principles about test impact, and assessment procedures which can influence significant educational decisions (evaluation of student learning) within a wider sociocultural context.” O’Loughlin’s (2013, p. 363) interpretation of LAL involves “the acquisition of a range of skills related to test production, test score interpretation and use and test evaluation in conjunction with the development of a critical understanding about the roles and functions of assessment within society.”

#### **1.4.2. Language Assessment Literacy Models**

Efforts have been made to establish what LAL should constitute for different stakeholders (Brindley, 2001; Davies, 2008; Fulcher, 2012; Inbar-Lourie, 2008; Pill &

Harding, 2013; Taylor, 2013). Brindley (2001, p. 129) argued that professional development programs designed to develop teachers' LAL competencies should emphasize curriculum-related assessment, utilize teachers' pre-existing knowledge, and provide expertise in components relevant to their particular needs, whether it pertains to detailed test development and analysis or the creation of informal classroom assessment instruments. Brindley's (2001) proposal of a professional development program model consisting of five components is outlined as follows:

- The social context of assessment
- The definition and description of proficiency
- Constructing and evaluating language tests
- The role of assessment in the language curriculum
- Putting assessment into practice

One of the core units in this modular program deals with the social, educational, and political dimensions of assessment. Different contexts have different standards and values for assessment. For instance, although formative assessment is given more weight than summative assessment in some nations, this may not be the case in others (Stabler-Havener, 2018). Ethical concerns, such as decision-making based on the validity or reliability of students' placement or admission tests, are also included within the social context. This refers to the issue of accountability, as the broader social setting influences stakeholders' perceptions of an assessment's validity. Therefore, teachers must take social considerations into account when preparing assessments and reporting their results (Stabler-Havener, 2018). Lastly, with regard to the political facets of assessment, teachers may minimize bias in assessment design by concentrating on the political exploitation of language assessments.

Defining and describing proficiency is another core unit in Brindley's (2001) model. The theoretical underpinnings of language tests and assessment procedures are addressed in this unit, including how to measure communicative competence and how to design assessment instruments in line with the discussions of Canale and Swain (1980) for communicative approaches to second language teaching methodologies. Furthermore, this unit involves test quality issues Bachman and Palmer (1996) put forth, such as the notion of "test usefulness," which emphasizes the interdependency and the necessity to establish a balance among six

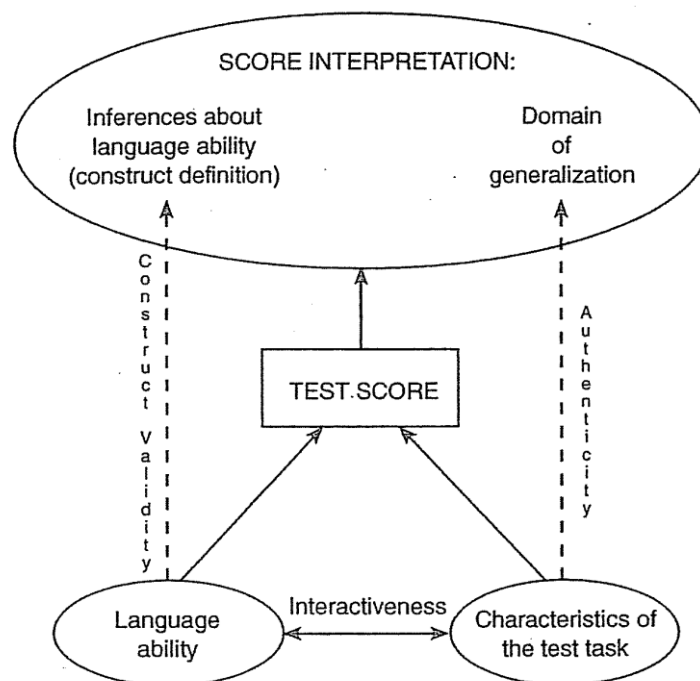
parameters: authenticity, reliability, construct validity, impact, and practicality (p. 41). Reliability is the congruency of test scores among the different domains of the test, while authenticity pertains to the relationship between test task properties and the properties of real-world tasks. Interactiveness measures how individuals' traits (language ability, background knowledge, and motivation) come into play during test-taking. Meanwhile, practicality is grounded in the administration of the test rather than test score interpretations (Fulcher & Davidson, 2007, p. 15).

The way in which Bachman and Palmer (1996) distinguish three aspects of language tests that had previously been considered to be components of validity, impact, authenticity, and interactiveness (Bachman, 1990) is another noteworthy takeaway from this discussion. An important notion that requires clarification here is "construct validity," defined as the degree to which a score can be interpreted as an indicator of the intended abilities or constructs for measurement (Bachman & Palmer, 1996, p. 21). To validate score interpretations, evidence must be presented demonstrating that assessment outcomes only reflect the components of language ability intended to be assessed. A "construct" specifies the ability that serves as a foundation for a particular test or a test task. In this regard, Bachman and Palmer (1996) emphasize the inefficiency of defining the construct of language ability for a given test situation within the scope of a limited framework involving one domain of language knowledge exclusively. This is particularly relevant as the construct in the target language use may encompass other facets of language knowledge, such as metacognitive strategies, topical knowledge, or affective responses (Bachman & Palmer, 1996, p. 41).

While acknowledging that impact is a component of the consequential basis for using tests and that authenticity and interactiveness are associated with construct validity, they argued that these elements hold significance to the development and implementation of language assessments and, therefore, deserve special attention. The term authenticity refers to the degree to which assessment tasks are suitable for the real-life use of the target language. The extent and type of involvement of the individual's characteristics, which includes language knowledge, metacognitive strategies, topical knowledge, and affective schemata, in completing a task, is referred to as interactiveness. The relationship between these components is visualized in Figure 1.

**Figure 1**

*The Relationship between Construct Validity, Authenticity, and Interactiveness*



*Note.* The Relationship between Construct Validity, Authenticity, and Interactiveness. From *Language Testing in Practice: Designing and Developing Useful Language Tests* (p.22), by L.F. Bachman, and A.S. Palmer, 1996, Oxford University Press. Copyright 1996 by Oxford University Press.

Embedded in this discussion is yet another key consideration, which is the impact of assessment and evaluation practices. Impact is outlined by Bachman and Palmer (1996) as the effects of assessments on society, educational systems, and all those involved. The impact of using an assessment on teaching and learning can be defined as the “washback effect” (Bachman & Damböck, 2018). Bailey (1996, p. 259) summarizes washback as the effects of testing on teaching and learning. The way teachers assess, including what and how they assess, can influence their instruction and how students learn, resulting in either positive or negative washback. Positive washback occurs when the use of assessment results in teachers and students engaging in activities that align with their beliefs about teaching and learning, while negative washback happens when an assessment leads to the use of teaching and learning activities that are not in line with these beliefs and teachers’ perceptions of what is important for students to learn. Large-scale high-stakes assessments, such as university admissions or standardized tests of language proficiency in various

contexts, often come after an extensive period of instruction, and the preparation for these assessments can subsequently dominate or overshadow other teaching and learning activities (Bachman & Damböck, 2018). When test content and methods do not align with classroom instruction, teachers may feel compelled to focus on the assessment content instead of the content they believe is important for students to learn. In the same vein, washback is an aspect of what Messick (1989) refers to as consequential validity, as it denotes how far the use of a test leads teachers and learners to take actions that enhance or hinder language learning in ways they would not do under normal circumstances, according to Messick (1996, p. 241).

The third unit in Brindley's (2001) professional development program model, constructing and evaluating language tests, is intended to equip teachers with test development and analytical skills. Statistical procedures like item analysis and item response theory are introduced through this unit. Subsequently, the unit on assessment in the language curriculum introduces criterion-referencing in language learning programs and methods for establishing criterion-referenced procedures to evaluate learners' progress and achievement. In this regard, the relationship between course objectives and assessment is emphasized. This unit examines defining objectives and outcomes as well as incorporating different alternative assessment techniques, such as observation schedules, portfolios, project work, and self—assessment techniques, into the curriculum (Brindley, 2001, p. 130).

Putting assessment into practice is the last component of Brindley's (2001) model. This unit calls for teachers to create an action plan for a workshop, enabling further inquiry and documentation of the topics introduced in the course, such as project proposals about test development, classroom research on assessment, or suggestions for policy development. While Brindley (2001) stressed the necessity of the first two units as critical, he suggested the following three units could be considered optional. Harding and Kremmel (2016), on the other hand, argued that professional development for language teachers should include all five aspects listed by Brindley (2001).

Inbar-Lourie (2008a) draws on Brindley's (2001) model to explore LAL competencies in terms of the following three dimensions of assessment knowledge: the rationale for



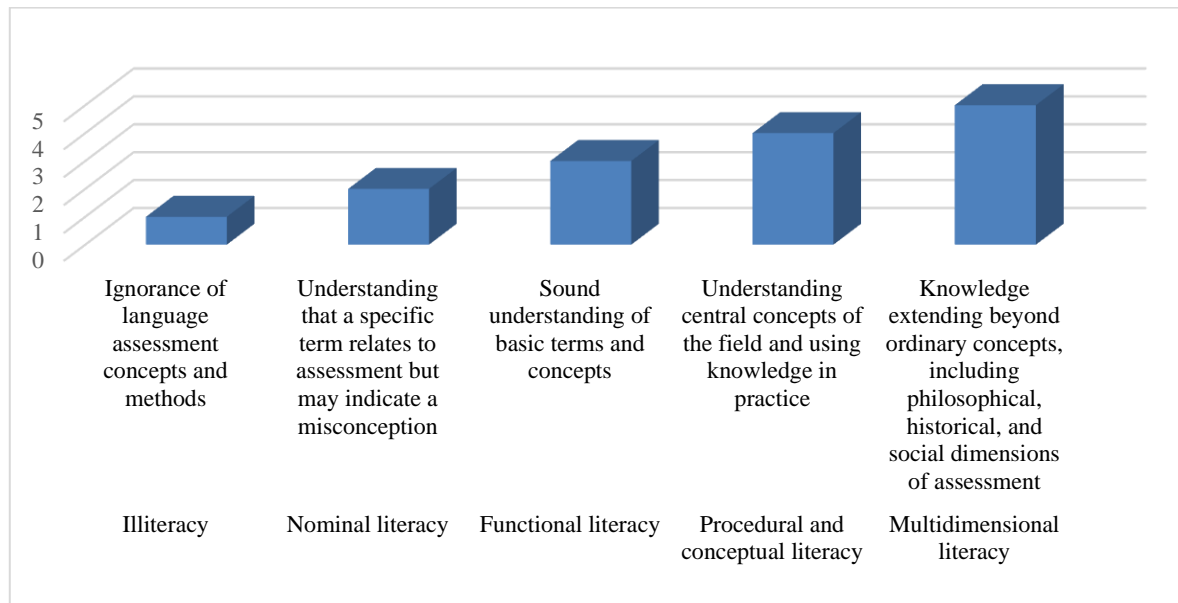
assessment (the “why”), the definition of the property to be assessed (the “what”), and the assessment process (the “how”) (p. 390). The first dimension of Inbar-Lourie’s (2008a) model emphasizes the rationale for assessment by considering the social views on knowledge construction and the place of language assessment in society. From this perspective, language assessment should be discussed as a social theme that can influence decision-making in several domains, such as civic, vocational, and educational. The second dimension described by Inbar-Lourie (2008a) corresponds to Brindley’s (2001) review of the theories and current perspectives underlying language assessment, such as understanding first language and culture’s role in second language acquisition. Moreover, issues related to ongoing discussions on the standards of English as an International Language (EIL) and the language competencies of multilingual speakers can be covered in this dimension. Expertise in LAL, therefore, necessitates the incorporation of these theories and approaches into assessment competencies. Stakeholders should be acquainted with recent developments in learning, teaching, and assessment of language features in order to design appropriate assessment instruments. Lastly, Inbar-Lourie (2008a) criticizes Brindley’s (2001) recommendation to offer the language assessment skills development modules “constructing and evaluating language tests,” which deals with test development and analysis, and “assessment in the language curriculum,” a criterion-referenced module, as optional. Since conducting sound assessments requires knowledge about both large-scale exams and classroom assessments, this division may cause participants to become knowledgeable in a certain area while remaining ignorant of other alternatives. While Brindley’s (2001) reasoning is rooted in the diversity of stakeholders’ settings, for instance, teachers engaging in classroom assessment primarily rather than developing tests and performing statistical analysis, Inbar-Lourie (2008a) contests the credibility of such policy. The fifth unit of Brindley’s (2001) model refers to staying informed about assessment initiatives and related research. Although Brindley (2001) notes this module is intended for assessment professionals only, Inbar-Lourie (2008a) suggests that due to the current improvements in language assessment applications, classroom teachers also assume an active role in these projects. To emphasize this argument, the “Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR)” (Council of Europe, 2001), which is intended to function as guidelines for explaining

qualifications, language learning objectives, and standards for achievement, is mentioned by Inbar-Lourie (2008a). The ability to act within this framework would entail an understanding of the levels and descriptors as well as competence in language teaching and assessment. Therefore, classroom practitioners must operate at the advanced level reserved for professionals by Brindley (2001) in order to adapt the framework to their needs. These discussions conclude that while LAL is built on the basis of assessment literacy knowledge, it has a unique intricacy of its own. Its focus on language warrants comprehension of assessment goals, characteristics, and methods that reflect the current theoretical perspectives in the field of language education.

Variations in stakeholders' expertise have led to an understanding that different stakeholders should exhibit different types of knowledge and degrees of competence (Pill & Harding, 2013, p. 383). Pill and Harding (2013) identified several groups of assessment stakeholders who would require different levels of LAL. To illustrate non-practitioners' LAL levels, Pill and Harding (2013, p. 383) developed a framework by adapting the stages of scientific literacy Kaiser and Willander (2005) previously presented. This framework offered a continuum-based approach to literacy and provided guidance for comprehending stakeholders' particular needs. The stages in the framework are detailed in Figure 2.

**Figure 2**

*Pill and Harding's (2013) LAL Stages and Descriptions*



*Note.* Pill and Harding's (2013) LAL Stages and Descriptions. [Adapted from] "Defining the language assessment literacy gap: Evidence from a parliamentary inquiry," by J. Pill and L. Harding, 2013, *Language Testing*, 30(3), 381–402. Copyright 2013 by Sage.

According to Pill and Harding (2013), "illiteracy" refers to a complete absence of knowledge or comprehension regarding language assessment concepts and methods, whereas "nominal literacy" denotes an understanding of specific terms connected to assessment, although accompanied by misconceptions. "Functional literacy" signifies a solid understanding of fundamental terms and concepts, and "procedural and conceptual literacy" indicates a grasp of central concepts and the ability to effectively apply this knowledge in practical settings. Finally, "multidimensional literacy" encompasses a larger scope of knowledge that surpasses ordinary concepts and involves an understanding of the "philosophical, historical, and social dimensions of assessment" (Pill & Harding, 2013, p. 384). Harding and Kremmel (2016) expressed their concerns about Pill and Harding's (2013) model due to its overemphasis on procedural and theoretical knowledge and failure to include social, ethical, and political competencies, with the exception of level four. Moreover, this model was unable to determine what level of LAL needs to be achieved for different stakeholders involved in assessment. Taylor's (2013) model of LAL addresses these issues by drawing on Pill and Harding's (2013) continuum model.

### 1.4.3. Taylor's (2013) Stakeholder Model for Classroom Teachers

Taylor's (2013) LAL model for stakeholders identifies four key stakeholders in assessment: "test writers," "classroom teachers," "university administrators," and "professional language testers," and outlines LAL components different stakeholder groups should possess under eight hypothesized dimensions involving the knowledge of:

- Theory
- Technical skills
- Principles and concepts
- Language pedagogy
- Sociocultural values
- Local practices
- Personal beliefs and attitudes
- Scores and decision-making

The first dimension, *Knowledge of Theory*, encompasses theories of language, language learning, and communicative competence, as previously discussed with reference to Canale and Swain (1980) in the models proposed by Brindley (2001) and Inbar-Lourie (2008a). Additionally, similar to Brindley's (2001) "defining and describing proficiency" module, this dimension includes assessment and testing theories such as validity, reliability, practicality, washback, and authenticity. The *Technical Skills* dimension covers test development methods and statistical expertise in item construction and analysis. Taylor's (2013) emphasis on the significance of this aspect for classroom teachers mirrors Inbar-Lourie's (2008a) argument about Brindley's (2001) professional development model, which provides test development and analysis or a criterion-referenced module as two distinct options.

The *Principles and Concepts* dimension, as Fulcher and Davidson (2007) describe, is related to the ethics of testing and fairness principles. The "concepts" branch, however, incorporates assessment theories previously laid out in the *Knowledge of Theory* dimension, which has led to criticism in the literature (Bøhn & Tsagari, 2021, p. 224). Meanwhile, the *Language Pedagogy* dimension addresses the interrelatedness of assessment, teaching, and learning. Pedagogical knowledge pertains to knowing how to transform disciplinary knowledge so that it becomes learnable for the students (Scarino, 2013, p. 316).

Pedagogical content knowledge, initially proposed by Shulman (1987), constitutes subject-specific professional expertise, including an awareness of factors facilitating or hindering mastery of particular subjects and knowledge of effective methods for expressing and framing target content so that it becomes understandable to learners. Taylor's (2013) dimension, therefore, encompasses the use of assessment to facilitate learning in addition to the delivery of content in a manner that fosters student understanding.

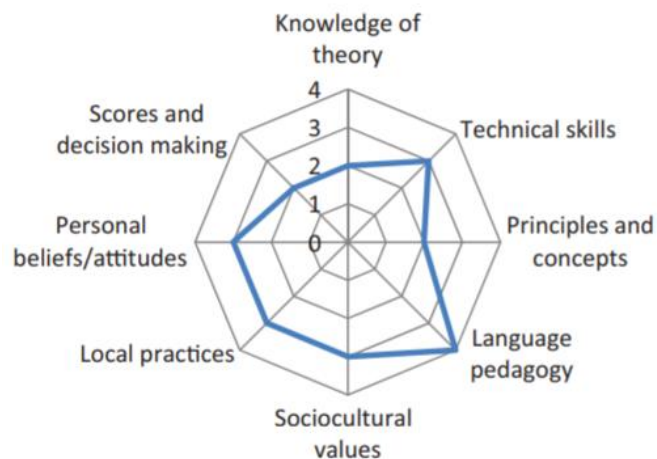
The *Sociocultural Values* dimension alludes to the influence of social and cultural values on assessment. This dimension is not extensively described by Taylor (2013); however, drawing on Inbar-Lourie's (2008a) model, it might be explained as the consideration of the social perspectives on the role of language assessment in society. Language assessment can be considered a social theme that can both influence and be influenced by various domains.

The dimension of *Local Practices* comprises curricula, policies, and legislation, as well as locally constructed assessment criteria, representing the institutional aspects of assessment in the school and classroom setting. *Personal Beliefs and Attitudes* reflect how teachers' perspectives, ideas, and assumptions affect their assessment practices. This entails judgments, biases, beliefs, conceptions, and outlooks about assessment that teachers bring to their practices (Scarino, 2013). Lastly, the *Scores and Decision-Making* dimension outlines awarding grades, categorizing students' performance, deciding whether students should be labeled as pass or fail, and making formative decisions to improve instructional practices (Bøhn & Tsagari, 2021).

Although Taylor examines different stakeholders and their roles in language assessment procedures, it should be noted that they are beyond the scope of this research. As aforementioned, this study centers around classroom teachers' LAL due to their direct involvement in assessment acts throughout the teaching and learning process. Therefore, the focus of investigation here is solely on the classroom practitioners. Figure 3 displays the LAL profile Taylor (2013) proposes for classroom teachers.

### Figure 3

*Taylor's (2013) LAL Profile for Classroom Teachers*



*Note.* This model was produced by Taylor in 2013, summarizing eight dimensions of LAL. From “Communicating the Theory, Practice and Principles of Language Testing to Test Stakeholders: Some Reflections,” by L. Taylor, 2013, *Language Testing*, 30(3), p. 410. Copyright 2013 by Sage.

According to Figure 3, classroom teachers are expected to have a strong understanding of language pedagogy while also having in-depth knowledge of *Sociocultural Values*, *Local Practices*, *Personal Beliefs and Attitudes*, and *Technical Skills*. However, they may exhibit lower comprehension in *Knowledge of Theory*, *Scores and Decision-Making*, and *Principles and Concepts*. In this regard, it can be inferred that classroom teachers are responsible for identifying and specifying language competence, understanding language assessment theories and methods, designing suitable assessment tasks, and accurately interpreting results while also considering local practices. Taylor’s (2013) LAL profile for classroom teachers is, therefore, concurrent with the consensus that classroom teachers should be required to possess all five components proposed by Brindley (Brindley, 2001; Harding & Kremmel, 2016; Inbar-Lourie, 2008).

As Harding and Kremmel (2016) pointed out, the dimensions of Taylor (2013) are speculative propositions. It could be argued that classroom teachers have a greater need for literacy in *Scores and Decision-Making* or *Personal Beliefs and Attitudes* than what Taylor (2013) has suggested. It also leaves the competencies that constitute its eight dimensions to interpretation since they are not clearly specified. Taylor’s LAL model provides a foundational framework for discussing and defining assessment concepts; however, due to

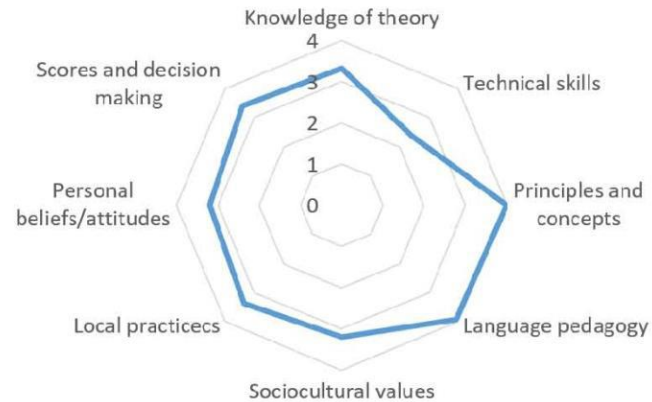
this hypothetical nature, the model lacks the level of detail and particularity that might be gained through further empirical research. As Stabler-Havener (2018) proposes, such investigations can shed light on the needs of different assessment stakeholders. Likewise, Bøhn and Tsagari (2021) suggest that evidence-based information is necessary to gain a better understanding of the context-sensitive aspects of teachers' LAL.

In light of the aforementioned framework, Bøhn and Tsagari (2021) conducted a study with five teacher educators to examine the relevancy of Taylor's (2013) LAL profile for classroom teachers in the Norwegian context. This study aimed to gather empirical data to operationalize and validate Taylor's (2013) hypothesized LAL dimensions. As part of this investigation, Bøhn and Tsagari interviewed Norwegian teacher educators to explore their perspectives on classroom EFL teachers' assessment needs. The findings highlighted that LAL is inherently a contextual concept, and investigating teacher educators' understanding of LAL in different contexts could be beneficial in revealing how their perspectives align with those in their study. Moreover, the results indicated that Taylor's (2013) model required adjustments to resolve commonalities in the existing dimensions and include new ones, such as "Collaboration Competence" and "Disciplinary Knowledge."

The emergent dimension, labeled "Disciplinary Knowledge," might be associated with Shulman's (1986) categorization of teacher knowledge, particularly the one pertaining to content knowledge. Content knowledge can be defined as an understanding of the subject matter and its organizational structures. According to Shulman (1986), comprehension of a subject for the purpose of teaching necessitates more than knowing its facts and concepts; it requires understanding how and why things are the way they are, for what reasons they can be rationalized, and in which conditions they might be challenged or rejected. Furthermore, Bøhn and Tsagari (2021) incorporate curriculum knowledge under the "Disciplinary Knowledge" dimension, which is another major point in Shulman's (1986) discussion. Curricular knowledge refers to programs designed to teach certain subjects at a given level. It describes the range of educational resources that are offered in connection with these programs. Figure 4 shows the LAL profiles of Norwegian classroom teachers.

**Figure 4**

*Classroom Teacher's LAL Profiles According to Norwegian Teacher Educators*

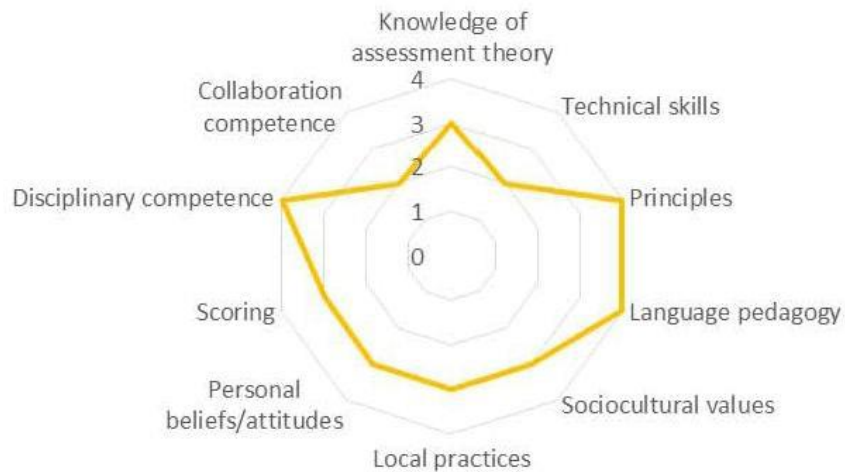


*Note.* From “Teacher Educators’ Conceptions of Language Assessment Literacy in Norway,” by H. Bøhn and D. Tsagari, 2021, *Journal of Language Teaching and Research*, 12(2), p. 228. Copyright 2021 by Academy Publication.

Bøhn and Tsagari’s (2021) revised model corresponding to Taylor’s (2013) dimensions based on the responses from teacher educators are visualized in Figure 5.

**Figure 5**

*Revised Version of Taylor’s (2013) LAL Profile based on Bøhn and Tsagari (2021)*



*Note.* From “Teacher Educators’ Conceptions of Language Assessment Literacy in Norway,” by H. Bøhn and D. Tsagari, 2021, *Journal of Language Teaching and Research*, 12(2), p.231. Copyright 2021 by Academy Publication.



#### **1.4.4. Research on Classroom Teachers' LAL**

In this section, the relevant studies on classroom teachers' LAL are presented.

Although academic discussions have stressed the importance of preparing teachers for their daily assessment practices, as a result of inadequate teacher training, only a small proportion of teachers feel ready to deal with the difficulties of classroom assessment (Stiggins, 2002, p. 762). According to Bachman (2000, p. 19–20), practitioners involved in the design and implementation of language assessments for educational purposes or applied linguistics research continue to undertake these tasks without adequate professional training. Alderson (2005, p. 4) also argues that teachers have insufficient assessment competence: “Tests made by teachers are often of poor quality, and the insight they could offer into achievement, progress, strengths and weaknesses is usually very limited indeed.”

Fulcher (2012) conducted an online survey to analyze language teachers' assessment training needs on a global scale. Out of 278 informants, 85% held a Master's or Ph.D. degree. The results of the study showed that language teachers required a textbook featuring activities that are appropriate for the language teachers who act as testers.

Coniam (2009, p. 227) reported major issues with the assessment tasks teachers design, such as inappropriate difficulty levels, assessment of unlearned content, and a failure to accurately reflect students' performance. Coniam (2009) investigated how an initial assessment course offered as part of the undergraduate English Language Teaching program affected the quality of tests teachers designed. Despite having the ability to improve their examinations, the final products still did not meet fundamental quality expectations. Furthermore, participants in this study reported facing limitations affecting their work standards, such as paucity of time, resources, and institutional cooperation.

A major factor in the effective use of assessment is ensuring that teachers are well-equipped with the necessary knowledge and skills by means of teacher training programs. Furthermore, Brindley (2001) suggested that teacher education programs should develop an understanding of how assessment practices align with the wider educational context. This involves identifying and dealing with the realities and limitations of assessment, defining and explaining language competencies, designing assessments in line with the language curriculum, and implementing them. Thus, Brindley (2001, p.129) proposes LTA training

targeting classroom practitioners should:

- focus on curriculum-related assessment
- exploit teachers' existing knowledge
- be adaptable to meet a wide variety of teacher needs

Exploring LAL from multiple perspectives may foster comprehensive frameworks that encompass its multifaceted nature and inform assessment training in educational settings. Vogt and Tsigari (2014) conducted a large-scale study with 853 participants across Europe to examine language teachers' LAL and training needs. They used a survey to determine foreign language teachers' educational backgrounds, institutional assessment responsibilities, and self-reported needs. Their study demonstrated that language teachers had insufficient assessment knowledge and required additional training in general. Notably, nearly 50% of the German informants reported that they had not received any training in LTA, a rate twice as high as those from Greece and Cyprus. This variation in LTA training among language instructors in different regions of the world confirms the complex nature of LAL and its multidimensional nature rooted in particular contexts. Vogt and Tsigari (2014) call for further inquiry in other educational settings to explore the contributing factors to variations in teacher LAL and help inform customized training programs that address the unique needs of classroom teachers.

O'Loughlin (2006) conducted research on a postgraduate elective TESOL course in Australia called "Assessment in the Language Classroom," which covered practical, theoretical, and social issues related to language testing. This study aimed to monitor participants' progress from the lecturer's perspective by analyzing students' contributions to the online forums. Particularly, two cases were examined in-depth. These cases represented the impact of standardized testing systems in Chinese universities as opposed to classroom-based assessment approaches in Australian primary education. Both students achieved the course's intended learning outcomes, but they had varying levels of readiness to accept new language assessment concepts due to cultural backgrounds and professional experience in assessment. Findings from O'Loughlin's (2006) study demonstrated the significance of embracing a learner-centric approach in the development and implementation of language assessment courses that can accommodate learners' diverse

cultural backgrounds and experiences.

In their exploratory study, Pill and Harding (2013) discovered that participants exhibited varying levels of knowledge gaps, misunderstandings, and uncertainties in language assessment. They tended to use imprecise terminology and lacked awareness of relevant professional expertise. These findings shed light on specific challenges in LAL, including fundamental misconceptions, knowledge gaps in testing responsibilities and processes, and a limited understanding of the constructs evaluated by specific language tests among non-specialists in the examined context.

In a syllabi review conducted by Jeong, only three non-LT instructors included topics related to statistics in their syllabi. Thus, the researcher identified two distinct instructor groups: those with a background in language testing and those without. Thus, the importance of language assessment and the dependence of its subject matter on the instructor's background has led to concerns about the quality of instruction provided by non-LTs, as there may be differences in the crucial topics that need to be covered in such courses (Jeong, 2013). She notes that the differences in the interpretation of assessment literacy can impact the courses' content, which varies significantly across six areas depending on the instructors' background, including test specifications, test theory, basic statistics, classroom assessment, rubric development, and test accommodation (Jeong, 2013, p. 355). Interview results confirmed non-LTs' weaker confidence in teaching technical assessment skills compared to LTs and their tendency to focus more on classroom assessment issues. This study highlights the need for high-quality teacher training in assessment to ensure a shared understanding of assessment literacy and consistent, quality course content in language assessment.

The purpose of Harding and Kremmel's (2019) study was to analyze LAL among stakeholders on a global scale. The survey, which was given to 1,529 stakeholders from 104 countries, including language instructors, test developers, and academics working in the field of language testing and assessment, had 71 items relating to 10 theorized elements. The results revealed different stakeholder profiles reflecting the variations in LAL across these groups. Language instructors, for example, ranked their expertise in test design lower than test developers or researchers. The Language Assessment Literacy Survey offered

empirical evidence supporting the perceived assumptions about stakeholders' LAL needs. Furthermore, the investigation validated the overall agreement between Taylor's (2013) LAL profiles and the data gathered. However, certain adjustments were found to be necessary in order to ensure congruity. Thus, refining existing LAL frameworks to optimize their pertinence and precision is emphasized by the researchers.

Recent studies on how teachers construct their LAL skills have revealed that such development reflects a dynamic and orderly pattern rather than a static accumulation of information (Crusan et al., 2016; Pill & Harding, 2013). Yan et al. (2020) conducted qualitative research in China to better understand how different stakeholders' LAL evolves. The researchers compared the LAL profiles of three groups of stakeholders: language testers, EFL teachers, and postgraduate students. Higher education language teachers were divided into two groups according to their educational background: non-English majors and English majors. According to the findings, non-English majors and secondary English teachers had higher LAL profiles. Moreover, the differences in teachers' and testers' profiles were found to be consistent with Taylor's (2013) stakeholder model. Yan et al. (2020) urge that more research be done in diverse contexts to see if these findings apply to other assessment systems and, if so, what might be the cause.

In the Turkish context, Hatipoğlu (2015) examined the LTA knowledge and needs of 124 pre-service EFL teachers. She reported that the majority of the participants had taken only one assessment course in pre-service education. In addition, more than half of the student teachers believed it was sufficient to aid their assessment duties in the future. The researcher argued that this perception might have been the reason behind in-service teachers' low LAL levels. Proving this point, more than half of the students were not aware of assessment-related topics and terminology. However, only a few participants addressed a need for training on innovative assessments, suggesting that they were not informed about recent developments. This result was also in line with Vogt and Tsagari's (2014) findings. Another compelling outcome of Hatipoğlu's study revealed that a substantial number of pre-service teachers requested assessment training for different age groups. This was most likely due to the absence of segregated teacher education and appointments in Turkey. We can conclude from these findings that the centralized teacher education programs and national teacher appointments in Turkey may not effectively cater to contextual varieties of

practitioners teaching at different educational levels.

From a different perspective, Mede and Atay (2017) examined tertiary-level EFL teachers' knowledge base, LTA needs, and attitudes in Turkish private universities. The researchers stated that the participants mainly lacked training in classroom-oriented assessment and thus required training in the following areas: informal assessment, self or peer assessment, and feedback. They also did not feel confident assessing productive, receptive, and integrated skills. On the other hand, they felt more confident in their knowledge of testing purposes, which was attributed to their familiarity with the placement tests in preparatory school. Moreover, most of the participants regarded their previous LTA training as unsatisfactory. Exploring what kind of LAL components they require may further enhance our understanding of its contextual dimensions.

In a more recent study, Gürsoy and Önalın (2020) surveyed 348 EFL teachers from Turkish primary and secondary schools to analyze their assessment perceptions, practices, and knowledge. Their study focused on how gender, experience, and the type of school environment influenced their perspectives and use of LTA. The findings showed that experienced teachers' assessment use was more diverse than that of novice teachers because the former group prioritized exploiting the feedback from their assessments to make decisions while constructing lessons and syllabi. In the same vein, experienced teachers valued student involvement in the assessment more than their novice counterparts. The reported LTA knowledge did not differ significantly between the two groups. These findings shed light on the complex relationship between teachers' experience, assessment practices, and knowledge. This research indicated that contextual factors, such as experience level and teaching environment, have an influence on teachers' assessment perceptions and practices.

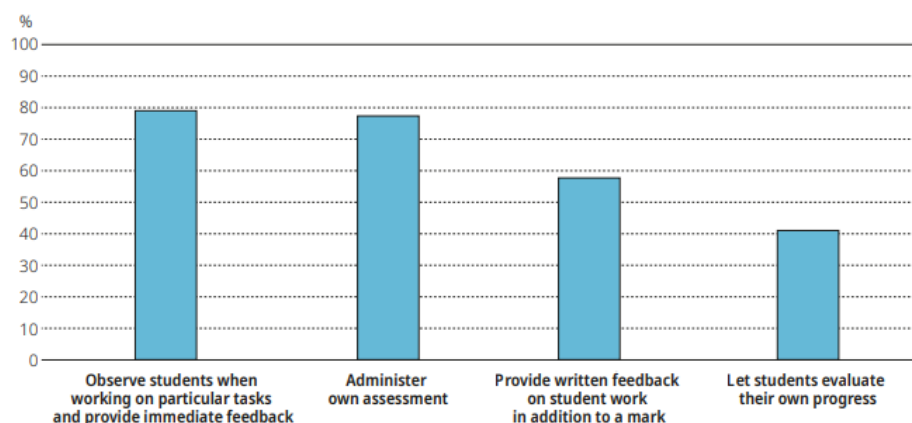
The above research findings emphasize the significance of well-designed initial teacher training and subsequent professional development programs in teachers' LAL. The effectiveness of LTA training is an indicator of how skilled teachers are in assessment. While there may be room for improvement in current training models in light of debates about dichotomies of formative and summative assessment, measurement against judgment, integrative and discrete assessment, and so on, it is equally important to acknowledge the

role that teacher educators have in building the foundation of pre-service teachers' assessment knowledge. Taylor's (2013) suggestions for early involvement in the design process of new assessments or assessment systems may also be pertinent when considering how immensely teacher education contributes to equipping language teachers with the skill set, knowledge, and principles they must acquire to construct an understanding of LAL. Therefore, evidence from research must serve as the basis for what should be planned or adjusted in assessment training to establish a LAL culture.

An examination of the Teaching and Learning International Survey (TALIS) reports reveals that the outcomes support the findings presented in the literature.

#### **1.4.5. Reports from the Teaching and Learning International Survey**

TALIS seeks to identify and address the challenges faced by teachers and school administrators in participating countries of the OECD by gathering information about what is going on in their educational institutions. TALIS 2018 asked teachers to report how frequently they use a set of four practices for assessing students in their classrooms (OECD, 2019). Two of the four assessment practices were more commonly used on average by teachers of OECD countries taking part in TALIS. 79% of teachers regularly assessed their students' progress by monitoring and giving them immediate feedback, and 77% administered their own assessments. Fewer teachers reported giving written feedback on coursework in addition to giving a grade (58%), while only 41% of teachers allowed students to assess their own progress, suggesting that formative assessment practices were less commonplace. Nevertheless, more teachers reported using assessment practices frequently in 2018 than in prior years' surveys (OECD, 2019). Figure 6 shows teachers' assessment practices in the participating OECD countries in 2018.

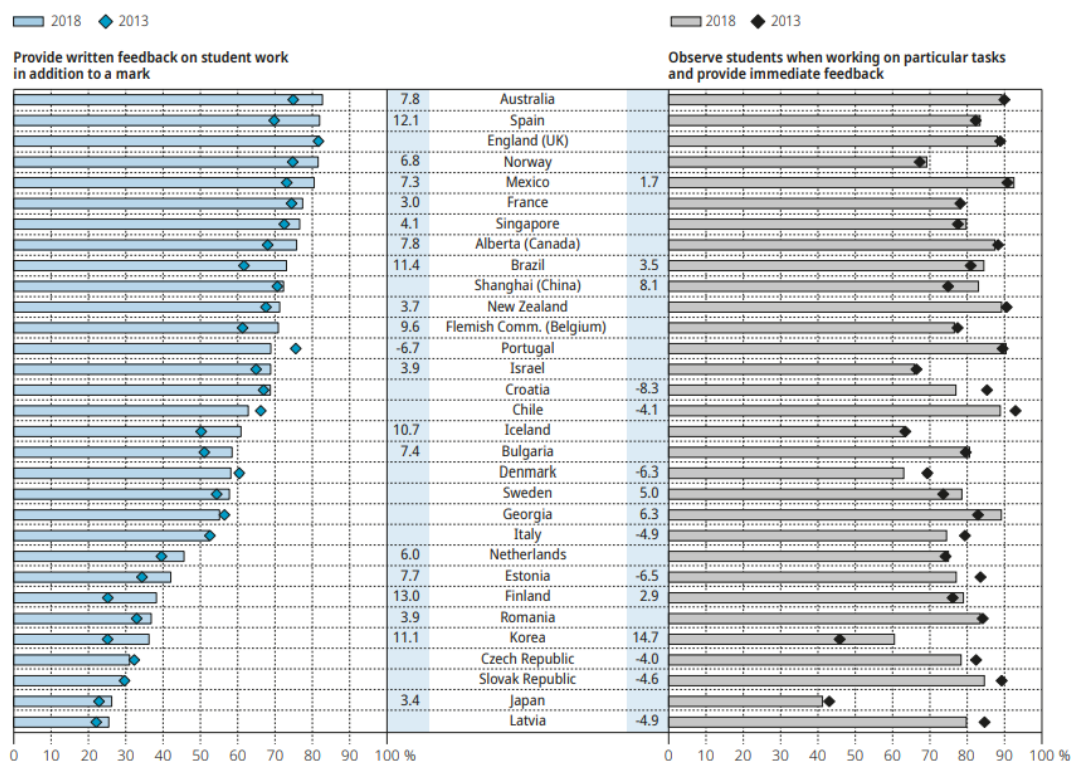
**Figure 6***Teachers' Assessment Practices in Participating OECD Countries*

*Note:* TALIS 2018 Results (Volume I): Teachers and School Leaders as Lifelong Learners, TALIS, Copyright 2019 by OECD Publishing.

When the change in teachers' assessment activities from 2013 to 2018 was analyzed, there were mixed results regarding their formative assessment practices. In terms of *“observing students on specific activities and providing them immediate feedback,”* eight countries and economies indicated a decrease, and seven showed an increase. Regarding *“actively involving students in their self-assessment,”* seven countries and economies indicated a decline, and ten others showed a rise. These results can be seen in Figure 7.

Figure 7

## Change in Teachers' Classroom Assessment Practices from 2013 to 2018



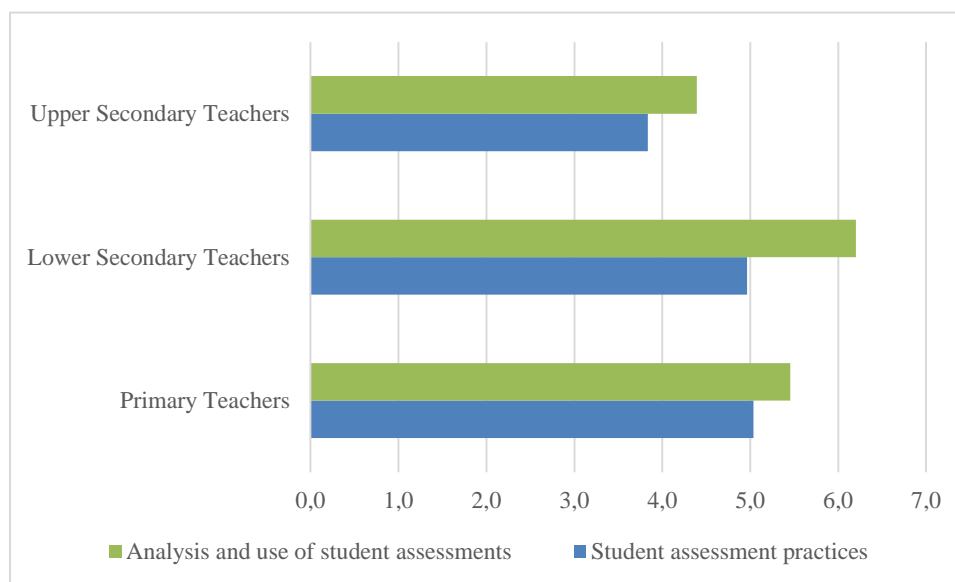
Note: TALIS 2018 Results (Volume I): Teachers and School Leaders as Lifelong Learners, TALIS, Copyright 2019 by OECD Publishing.

In the case of Türkiye, the TALIS 2018 reports indicate that Turkish upper secondary teachers had the lowest need for professional development (3.8%) in “*student assessment practices*” of all participating OECD countries (OECD, 2019). Similarly, Turkish upper secondary teachers ranked second lowest (4.49%) for the professional development demands in “*analysis and use of student assessments.*” The reports also suggested a decrease in teachers’ professional development needs between 2008 and 2018. Notably, 9.2% of Turkish lower secondary teachers acknowledged the need for training on student assessment in 2008. By 2018, this number had dropped to 5.0%, representing a -4.3% decrease. Figure 8 illustrates the TALIS 2018 percentages of teachers, indicating a high need for professional development across all three levels of education in Türkiye.



**Figure 8**

*Teachers Reporting a High Level of Need for Professional Development in Türkiye*



*Note:* TALIS 2018 Results (Volume I): Teachers and School Leaders as Lifelong Learners, TALIS, Copyright 2019 by OECD Publishing.

Moreover, there was an increase in the number of teachers in OECD countries who reported frequently using student assessment practices in their instruction by 2018. The percentages of teachers who reported that they “frequently” or “always” use the following assessment methods in their class, according to TALIS 2018 data, can be seen in Table 1.

**Table 1***Teachers' Classroom Assessment Practices from TALIS 2018 Reports*

Teaching Experience	Administer own assessment		Provide written feedback on student work in addition to a mark		Let students evaluate their own progress		Observe students when working on particular tasks and provide immediate feedback	
	%	S.E.	%	S.E.	%	S.E.	%	S.E.
Primary Teachers	83,1	1,0	65,0	1,2	68,0	1,6	90,7	0,9
Lower Secondary Teachers	83,0	1,0	54,7	1,0	59,4	1,1	84,5	0,8
Upper Secondary Teachers	80,3	0,8	53,1	1,0	60,1	1,1	81,5	0,8

When Talis reports are examined, it becomes evident that assessment remains an issue that requires improvement for teachers at both international and national scales. It is clear that many countries are still slow to adopt alternative assessment methods and to keep track of the global movement towards formative assessment. Nevertheless, the change that Türkiye has undergone over the last 10 years indicates that assessment has taken an important spot in the educational agenda of the country. In providing teachers with the necessary assessment skills, it has arguably become increasingly important to follow the developments in this area and to promote a culture of assessment, taking into account the specific contextual circumstances of each country. Evidence from these data also supports the view that LAL is a context-dependent, complex, and unique concept.

## 2. METHODOLOGY

This study aimed to understand the conceptions of LAL among teacher educators within the Turkish context using a qualitative method of inquiry. The rationale behind choosing a qualitative approach was due to its flexibility and exploratory nature, facilitating the production of in-depth and nuanced insights. This approach is effective in exploring complex phenomena in real-life contexts. It offers an interpretative, naturalistic perspective on the social world, allowing the researchers to gain in-depth and contextually rich data.

### 2.1. Research Design

The conceptualization of LAL is heavily influenced by contextual factors, personal beliefs, attitudes, and ideologies. Therefore, this research employed a qualitative method due to its strength in enabling the investigation of complex phenomena by examining individuals' perceptions, insights, and experiences (Braun & Clarke, 2006). In an attempt to gain an understanding of informants' perspectives and the meanings they attach to LAL, this study employed a phenomenological approach (Moustakas, 1994). Phenomenological design is commonly utilized in studies with small sample sizes that investigate participants' perspectives on a subject, event, or their skills, abilities, and attitudes in-depth. Furthermore, the phenomenological approach is suitable for studying the experiences and perceptions that shape one's understanding of LAL.

### 2.2. Participants

The present study included five teacher educators who had been selected from different universities across Türkiye. The selection of this group was guided by the principles of purposive sampling, a non-probability sampling technique that is commonly utilized in qualitative research (Creswell, 2013). This sampling method is used when a researcher intends to collect detailed information from a specific subset of a population, providing 'information-rich' participants who could offer valuable insights and in-depth information that is pertinent to the research questions and context selected (Patton, 1990, p. 196).

The teacher educators chosen for this study were between the ages of 34 and 48, with diverse experience and expertise in the field of ELT. Four of the five teacher educators held Ph.D. degrees in ELT. The fifth participant was a Ph.D. holder in English Language and Literature (ELL). Participants' academic and professional backgrounds were anticipated to

contribute to a variety of insights and perspectives. Their areas of expertise ranged from intercultural communication, intercultural competence, multilingualism, the use of technology in language education, individual differences, academic writing and feedback, reading comprehension, and academic vocabulary to ELL and American and British Literature (ABL). Although all participants had some degree of involvement in implementing classroom assessments, only two of them specialized in language assessment with publications on the subject.

The educators' duration of employment in the faculty varied from relatively new to more experienced, with an average tenure of 18 years. Similarly, the length of their teaching experience in the field of ELT also varied, extending from 5 to 18 years. It was anticipated that the TEs' professional histories would reflect the variety of professional experiences needed to explore the phenomena under investigation fully. In consideration of the ethical responsibilities and participant confidentiality, pseudonyms were employed in the form of "Teacher Educator 1" (TE1), "Teacher Educator 2" (TE2), and so forth to respect and maintain the anonymity of their identities. The demographic information of each participant is detailed in Table 2.

**Table 2***Demographic Information of the Study Group*

Participant	Age	Location	Area of Expertise	Years in Teaching	Teaching Experience	Assessment Experience	Assessment Publications
TE1	36	Hatay	Intercultural communication, intercultural competence, multilingualism, technology in language education	10 (Lecturer), 4 (Teacher Educator)	Culture, Cultural Communication, Material Development and Adaptation	Ten years in modular course-based education in a preparatory school	None
TE2	34	Afyon	English language teaching, Language Assessment and Evaluation, Individual differences	10 (Includes private school and university)	In-service training for English teachers focusing on Assessment and evaluation	Testing Office at the School of Foreign Languages, proficiency exams preparation	Doctoral thesis on language assessment
TE3	41	Isparta	English Language Teaching (ELT), Language Assessment, Test anxiety, Assessment literacy	18 (ELT Department)	Language Assessment and Testing courses in the ELT department	Master's level language assessments, focus on reading, grammar, vocabulary	Publications on primary-level English teachers' assessment beliefs and practices, Assessment literacy
TE4	35	Afyon	English language education, Academic writing, Reading comprehension, Academic vocabulary	5 (Faculty member in English language education department)	General assessment and evaluation techniques	1.5 years as a lecturer in a preparatory school, exam preparation and evaluation	None
TE5	48	Nevşehir	English Language and Literature, American and British Literature	7 (Teacher educator), 15 (Preparatory school)	Comprehensive assessment and evaluation courses	Preparation for YDS (English Proficiency Exam for Academic Purposes)	None

**2.3. Data Collection Tools**

Data collection was conducted using a semi-structured interview guide and closed-response questionnaire items developed by Bøhn and Tsagari (2021). The interview guide focused on informants' comprehension of LAL, addressing the first research question.

Open-ended questions were structured to gather in-depth reflections on teacher LAL. Additionally, follow-up questions were included to encourage elaboration and further exploration of previously discussed ideas.

The interview guide was translated into Turkish by the researcher, with feedback from an English-proficient expert, to ensure translation accuracy. Before the main data collection phase, a pilot interview was conducted with a teacher educator specializing in both ELT and Curriculum and Instruction programs. This pilot interview lasted for approximately 50 minutes, and the clarity and understandability of the questions were thoroughly discussed with the expert following the session. As a result of this discussion, several revisions were made to increase the clarity of certain questions.

Care was taken to make only a small number of modifications to fit the context of Türkiye. Specifically, references to national tests such as the English Proficiency Exam for Academic Purposes (YDS) and Higher Education Institutions Foreign Language Test (YÖKDİL), the Education Act, and The Basic Law of National Education, were incorporated into Items 20 and 25. Furthermore, the wording of Item 22, initially phrased as “it is important to have knowledge of languages,” was adjusted to more accurately reflect the intended meaning, becoming “it is important to have knowledge of languages and language learning.” Furthermore, Item 23 was supplemented with a follow-up question, specifically, “What benefits does knowledge of the CEFR provide to the teacher?” in order to provoke a more in-depth response. These adjustments were implemented to ensure that the interview guide was as comprehensive, relevant, and effective as possible.

As for the second research question of this study, the relevancy of Taylor’s (2013) LAL profile for classroom teachers from Turkish TEs’ perspectives was explored. To this end, a six-point Likert scale questionnaire was adopted from Bøhn and Tsagari’s (2021) study, which was structured considering the implications of Taylor (2013). Table 3 presents a detailed description of the items and corresponding dimensions in the questionnaire, taken directly from Bøhn and Tsagari (2021).

**Table 3**  
*Operationalization of Taylor's LAL Dimensions*

<b>Dimension</b>	<b>Item No</b>	<b>Item</b>
Theory	22a	It is important to have knowledge of languages and language learning
	22b	It is important to have knowledge of theories of communicative competence.
	27	It is important to have knowledge of assessment/testing theory (issues such as «validity», «reliability,» etc.).
Technical skills	15a	It is important to have knowledge of how language testing can be used appropriately in the classroom.
	15b	It is important to have knowledge of how good items are created.
	15c	It is important to have knowledge of statistical measures in order to interpret results from multiple-choice tests (e.g., mean, standard deviation, measurement error)
Principles and concepts	28	It is important to have knowledge of ethical issues (fairness, use of assessment results for purposes other than what was intended, etc.).
Language pedagogy	12	It is important to have knowledge of how assessment can promote learning.
	13	It is important to have knowledge of self-assessment.
	14	It is important to have knowledge of peer-assessment.
	16a	It is important to have knowledge of formative assessment.
	18	It is important to have knowledge of how assessment can be used as a diagnostic tool.
	21	It is important to have knowledge of how assessment can be used to motivate students.
	30	It is important to have knowledge of how to communicate assessment results in appropriate ways (e.g., how to explain the results from national tests).
Sociocultural values	31	It is important to have knowledge of how values in society may affect assessment
Local practices	24	It is important to have knowledge of the national curriculum (the general part and the English subject curriculum).
	25	It is important to have knowledge of the Regulations of the Education Act.
	26	It is important to have knowledge of local issues that may be relevant for assessment (e.g., the local educational authorities' assessment criteria, how teachers in your local district assess student performance, etc.)
Personal beliefs/attitudes	32	It is important to have knowledge of one's own view on assessment (values, outlook) may affect the assessment.
Scores and decision-making	16b	It is important to have knowledge of summative assessment.
	29	It is important to have knowledge of how to use rating scales/scoring rubrics.

*Note.* Reprinted from “Teacher Educators’ Conceptions of Language Assessment Literacy in Norway,” by H. Bøhn and D. Tsagari, 2021, *Journal of Language Teaching and Research*, 12(2), p. 222-233. Copyright 2021 by Academy Publication.

## 2.4. Data Collection

Data were collected through interviews carried out in Turkish, the participants’ native language, according to their preference and to facilitate ease of communication and ensure that their perspectives were accurately captured. Interviews were conducted either face-to-face ( $n=3$ ) or via Zoom ( $n=2$ ), depending on participant preference and circumstance. The duration of these interviews ranged from 40 minutes to two hours. Details of interviews conducted with teacher educators are presented in Table 4.

**Table 4**

*Duration and Mode of the Interviews*

<b>Participant</b>	<b>Interview Duration</b>	<b>Mode</b>
TE1	2 hours	Zoom
TE2	48 minutes	Face-to-Face
TE3	50 minutes	Face-to-Face
TE4	1 hour 22 minutes	Face-to-Face
TE5	40 minutes	Zoom

Prior to the interview, each participant was informed about the purpose of the study. Their consent was obtained, covering the audio recording of the interview, its transcription for subsequent analysis, and the anonymous use of their data for publication. These measures were taken to ensure compliance with the ethical considerations inherent to qualitative research.

Each interview was audio-recorded, following ethical protocols and with explicit participant consent, then transcribed verbatim. The verbatim transcriptions were translated into English to ease the analysis process. In instances where further clarification or additional insight was needed, a series of follow-up correspondences were conducted via email post-interview.

## 2.5. Data Analysis

The open-ended responses were analyzed using thematic analysis in four stages. First,



the interview transcripts were carefully examined and reread multiple times in order to become familiarized with the content. Then, tentative codes were generated by revising the existing literature, particularly focusing on Taylor (2013) and Bøhn and Tsagari (2021). These initial codes served as a framework for exploring the emergent patterns in the data, and further adjustments and revisions were expected. Once a portion of the data had been coded and recurring concepts became evident, the codes were grouped under broader themes. At this stage, the source papers were re-examined multiple times to ensure that the identified themes effectively represented the coded data. Each phase of the analysis was meticulously monitored and guided by the thesis supervisor.

Table 5 provides an overview of the list of preliminary codes and overarching themes identified through the analysis of literature review and interview data.

**Table 5***Overview of Codes and Themes Incorporating Taylor's (2013) LAL Dimensions*

<b>Main Theme</b>	<b>Theme</b>	<b>Code</b>	<b>Explanation</b>
Theoretical Knowledge of Assessment	Assessment-specific Knowledge	Assessment philosophy, concepts, purposes, techniques, and design Alternative/traditional assessment Selecting assessment tools Test theories Validity Reliability	This category combines Taylor's (2013) 'Knowledge of Theory,' 'Technical Skills,' and 'Principles and Concepts' dimensions
Competence in Scoring and Statistics	Assessment Principles	Washback effect Fairness	
	Technical Skills	Statistics Test analysis Scoring/grading related to numbers and statistics Discriminating students Distribution of scores	This category involves knowledge of statistical measurements in Taylor's (2013) 'Technical Skills' dimension and scoring in the 'Scores and Decision-Making' dimension
Language Pedagogy Knowledge	Pedagogical Knowledge	Monitoring students' development Making decisions about teaching and learning Giving feedback Awareness of individual differences	This category involves knowledge of pedagogical content in Taylor's (2013) 'Language Pedagogy' dimension and issues related to decision-making from 'Scores and Decision-Making' dimension
Content Knowledge	Disciplinary Knowledge	Grammatical structures Vocabulary Language skills (speaking, writing, reading, and so on.) Language acquisition Language learning theories Communicative competence Constructivism Sociocultural theory Knowledge of the curriculum	This category involves issues related to disciplinary knowledge, such as knowledge of language learning theories, approaches, and methods such as SLA, communicative competence, constructivist theory, CEFR, Standardized English, English as a world language, English as a lingua franca
Social Context Knowledge	Social Context	National education system National assessment system Systemic demands Institutional demands Experience, familiarity, and preconceptions Students' readiness and preparedness Academic expectations Job satisfaction	This category combines issues from Taylor's (2013) 'Sociocultural values,' 'Local practices,' and 'Personal Beliefs and Attitudes' dimensions. It operates as an all-encompassing label for components related to institutional demands, collective assessment beliefs, and societal attitudes toward assessment, as well as teachers' personal beliefs and attitudes
	Personal Beliefs and Attitudes		

The quantitative data analysis was conducted using a statistical package for the social sciences software. Due to the small sample size, median and mean scores were computed to demonstrate which questionnaire items received the highest and lowest scores. The responses to the items were measured using a six-point Likert scale that ranged from “Not at all important” to “Very Important.” The findings obtained from a total of 28 items (seven items in the dimension of *Language Pedagogy*, three items each for the dimensions of *Knowledge of Theory*, *Technical Skills*, and *Local Practices*, two items in *Scores and Decision-Making*, and one for the dimensions of *Principles and Concepts*, *Sociocultural Values*, and *Personal Beliefs and Attitudes*) were presented under the findings section.

### 3. FINDINGS

This section presents the findings from the qualitative and quantitative data analysis in line with the research questions:

1. How do Turkish teacher educators conceptualize LAL?
2. What relevance does Taylor's (2013) LAL model have for Turkish teacher educators?

The findings obtained from the responses to interview questions were presented with supporting evidence obtained from questionnaire responses in order to show the degree of importance teacher educators attribute to the LAL dimensions Taylor (2013) outlined.

#### 3.1. Teacher Educators' Conceptions on LAL

The first research question of this study was formulated as "How do Turkish teacher educators conceptualize Language Assessment Literacy?".

The investigation of the responses yielded the following themes:

##### *Assessment-specific Knowledge*

The theme of *Assessment-specific Knowledge* addresses several issues, including assessment concepts, the purpose of assessment, and test theories. Informants generally agreed on the significance of understanding the philosophy of assessment, which entails why we carry out assessment, what is being assessed, how it is being assessed, and what can be accomplished with the data obtained from assessment. The TEs suggested that classroom teachers who are well-informed about the objectives, procedures, and theory of assessment would, in turn, be capable of carrying out assessment-related tasks in their school settings. The following interview excerpt reflects how TE4 discussed this notion in detail:

*I believe that teachers, first and foremost, should understand the philosophy of why they are conducting assessments, why they are measuring student learning or performance, and what they will do with it afterward. I think if they understand the philosophy of this, that is, if they understand and examine the purpose of assessment well, they will have a better idea about the assessment techniques they need to use, whether it is the measurement techniques, tools, or methods. Therefore, I expect them to understand the rationale of the assessment first.*

Meanwhile, TE3 pointed out that teachers need to be able to distinguish between different assessment purposes and understand the reasoning behind the selection of a particular method in a particular situation. Moreover, he stressed the necessity of establishing the validity of the assessment by explaining that each item on an exam should represent what it is intended to measure:

*First, the teacher needs to know what assessment and evaluation are. For them, it is usually just about the ten or fifteen matching or multiple-choice questions they write on exams. They need to separate these theoretical concepts from each other. They should know why they are conducting these assessments because every assessment should have a purpose. Each question they ask should be measuring something; they need to be aware of this. If they are aware of this, the rest follows naturally.*

Furthermore, TE3 underlined a lack of awareness among test developers and teachers alike in terms of how to align assessment methods with the nature of language skills they are aiming to assess. The following excerpt clearly reflects the interconnected nature of LAL components, particularly subject matter expertise and assessment-specific competence:

*Of course, not every question format is suitable for every skill. Typically, we prefer multiple-choice questions for reading skills, but we need to break this habit. [...] So, the type of assessment for each skill is different. We discussed this for speaking as well. Some instructors assess speaking with writing. Our YDS exams have questions like "Select the expression that fits this dialogue," which does not involve production. Speaking is a form of production, but we try to make students find it not by producing it but by selecting it in a national standardized test.*

TE5, on the other hand, explained the relationship between teaching and assessment as follows:

*I do not separate assessment from teaching... Do teachers not know different teaching methods now? Aren't they supposed to know? In that case, they must also know the types of assessment. Because English is not like a math lesson; it has at least four skills. Would assessing all four skills in the same way yield reliable results? Therefore, they are obliged to know different assessment types, such as process assessment and alternative assessment. They must know all of these for sure.*

TE5 suggested that just as teachers are responsible for being knowledgeable about different teaching methods, they are also expected to have an understanding of various assessment types. He stressed that English language instruction is different from other

teaching subjects due to the multifaceted nature of languages. Therefore, language assessment must reflect this nature. This idea also supports what TE3 discussed about the importance of tailoring assessment to the specific skills being evaluated. They both argue that using the same assessment method or approach for different situations may fail to produce reliable outcomes. Therefore, teachers are expected to be well-versed in different kinds of assessment, such as formative assessment and alternative assessments. TE2 expanded on this issue by expressing that teachers should keep up with the recent methods and trends in assessment, as presented in the following quote:

*In my opinion, teachers should certainly be proficient in the latest techniques related to assessment and evaluation [...] This includes keeping track of innovations like individualized learning or e-portfolios. Teachers should also adapt to these innovations.*

TE3 supported this notion when he discussed why teaching alternative assessments is their priority in pre-service teacher education programs. They highlighted the shortcomings of traditional assessment techniques in evaluating students' abilities. Therefore, they were actively making an effort to train prospective teachers on recent methods and techniques in assessment.

*They graduate with sufficient theoretical background, but the theories we teach in pre-service teacher education require more effort; that is, we focus more on alternative assessments. We believe that students' skills cannot be measured by written exams and multiple-choice tests, so we teach them accordingly...*

While elaborating on the issue of high-stakes tests, both TE1 and TE3 referred to exams like TOEFL and IELTS. The following excerpts are taken from the interviews conducted with TE1 and TE3, respectively:

*For example, YDS mainly concentrates on reading, whereas it does not measure any productive skills. So, this exam is supposed to assess skills based on production more. There is already groundwork being laid on these matters. In a few years, this exam is anticipated to have a format that will measure various skills. We aim to imitate how TOEFL or IELTS is handling this issue. Language education is not an arena in which Turkey dominates in the world. We need to examine what Western societies are practicing in this field and work on what is possible in our conditions in Turkey.*

*Because these tests [large-scale international exams] use different assessment methods. YDS and YÖKDİL, well, they are somewhat similar, but exams like TOEFL and IELTS use assessment methods that we do not generally use in our context. I mean, they assess*

*speaking and writing. In reading, they do not just give the student a paragraph and ask typical questions as we do; as I mentioned earlier, they provide a picture, tables, or graphs and ask questions from within the visuals. So, it is good to know different methods.*

Although both TEs acknowledged that classroom teachers should essentially focus on the properties of classroom assessment rather than large-scale high-stakes testing, they believed that teachers should still be aware of the question types and evaluation processes. Even if these types of tests do not transfer into classroom settings, teacher educators still find them to be valuable for teachers in terms of providing examples assessing different skills, particularly speaking and writing.

Overall, the main conclusions that can be drawn from the TEs' discussion on the *Assessment-specific Knowledge* theme is an emphasis on the practical elements of assessment, including how to select appropriate assessment methods for specific situations, how to design assessment tasks in line with the assessment purposes, and how to prepare good items. Therefore, equipping pre-service teachers with a solid background in language assessment theories, including their driving philosophies, current techniques and strategies, and a collection of alternative resources, appears to be the initial stage of LAL development.

### ***Assessment Principles***

Taylor's (2013) *Principles and Concepts* dimension has been revised in light of Bøhn and Tsagari's (2021) discussion. Thus, the theme of *Assessment Principles* was used to create a distinction between the knowledge of assessment theories, concepts, and principles. Principles were reserved for concerns of ethics, fairness, and washback effect. The most prominent issue that surfaced from the interviews was the washback effect.

The washback effect has been suggested to be unfavorable or beneficial for learning since it can either encourage or hinder the fulfillment of desired outcomes. Washback, in the words of Alderson and Wall (1993), describes the driving force that leads both teachers and students to engage in things "they would not necessarily do otherwise" (p. 117). TE5 clarified the relevance of the washback effect in classroom assessment as follows:

*Whatever we are assessing, we are implicitly telling the students that they should learn this. Therefore, the things we measure in the teaching and learning process are more*

*valuable for the students, and they tend to focus more on them. It is like this: The reason why children in Turkey cannot speak English is a question that always comes up. And the problem is that we never assess the speaking ability of the students. This is the washback effect. Students think that since this is not part of the examination, there is no need for them to learn it, so they focus more on grammatical structures and vocabulary.*

Similarly, the washback effect in the case described by TE4 demonstrates how assessment practices can influence students' behavior and attitudes toward learning. This suggests that an overemphasis on result-oriented assessment in education systems can cause students to concentrate on obtaining high scores instead of remaining engaged with the subject matter afterward. The related quotation from TE4's discussion is provided below:

*You have announced the final marks, so even if you give them feedback, the student decides, "I'm done with this course. I do not need to learn more about it," and does not study for that course again. Because the education system in Turkey is a bit more product-oriented and result-oriented, they say my academic transcript is more important.*

TE3 drew attention to why leveraging assessment as an incentive for learners and for pointing out their strengths, instead of using it as a penalty, is important in classroom assessment as reported under:

*This is not the perfect analogy, but it is an effective tool because if you say, "This topic will be on the exam," the students' motivation to study that subject becomes different. If you say it will not be on the exam, their motivation to study will be different. Therefore, assessment is very important and contributes significantly to the teaching process.*

TE2 used a similar expression, "weapon," to describe this concept:

*This topic is very important because assessment can either significantly demotivate students or be used to motivate them. It is crucial that teachers use it to motivate students and show them their strengths instead of using it as a weapon. So, what are your strengths? What else can we do in the future? A balance is essential, I believe.*

Finally, regarding the issue of fairness, TE3 reported the relevance of ethical aspects of assessment as follows:

*If you conduct incorrect assessments and evaluations, there can be many consequences, and you can affect the students' lives. You might end up giving unfair grades, so this issue is very important.*



### ***Technical Skills***

The theme of *Technical Skills* covers the statistical elements involved in assessment rather than the practical components of assessment. This theme, which was explicitly drawn from Taylor's (2013) LAL dimensions, provided comparability between the working conceptualization of this dimension and the quantitative findings of the current study. TEs believe that classroom teachers rarely have use for complex statistical measurements for their routine instructional activities. Therefore, they suggested that the practical elements of assessment, such as understanding the educational philosophies of assessment or skills to prepare effective assessment tasks, take priority over statistical know-how in the context of classroom-based assessment. TE4 put this into perspective as the following excerpt shows:

*I find that there are many kinds of statistics to be used in assessment, such as t-scores, z-scores, and so on, but in general, I observe that teachers do not need these statistics in school settings. In this sense, I think assessment philosophy and then the use of assessment tools appropriate for these philosophies, the preparation of questions, and item analysis seem to be more important.*

In agreement with TE4, TE1 referred to the rationale behind the need for complex statistical calculations in assessment. He argued that larger class sizes might rely on more statistical data to evaluate overall student performance. However, he also mentioned as the class sizes in MoNE have been decreasing in recent years, there is no longer a requirement for classroom teachers to be competent in statistics:

*[...] It will depend on the group you are teaching. How many students do you have in your class? In the past, the classrooms were very crowded in Turkey. Twenty years ago, class sizes were up to 40 students. But now the average class size consists of about 25 students. The statistical data is important, of course. Depending on the number of students, what percentage of the class has passed and what percentage have failed? What is the average? So, this average provides us with the information that the class is successful in this subject. This allows us to reflect more accurately. But if you are teaching a smaller group, you need to monitor on an individual basis, such as the portfolio assessment. Our graduates usually work at the Ministry of National Education, and their systems automatically generate the statistical data for teachers. However, if we are talking about drawing conclusions from the numbers, more training can be provided on interpreting what the lowest and the highest scores in the class indicate.*

Instead of concentrating on mathematical calculations, as TE1 offered, teachers should be adept in understanding the meaning of these numbers and what they imply for the

students in the context of the classroom, which is, in fact, a concern of using assessment data to support decision-making processes rather than technical competences. TE4's opinions regarding this issue support the distinction between statistical components and aspects of decision-making.

*But I think the main issue to be considered is, first of all, to make a first overall evaluation of the class, whether there are many underachieving students, whether there are high achieving students that require more support, what I can do in supplement or what kind of precautions I can take for the underachieving students. Secondly, in addition to the characteristics of the student, it is enough for me to know these statistics to get a rough idea of the validity or reliability of the exam I administered and whether my exam questions were extremely difficult. No one could pass whether they were sufficiently discriminating, how the distribution was, and so on.*

As TE4 suggests, gaining insights into the effectiveness and quality of the assessment tasks is associated with how assessment data can inform teachers. In addition, TE3 has expressed that statistics in assessment was one of his areas of expertise. His perspective on this matter is as follows:

*These elements are surely important assessment concepts, but they might not significantly affect classroom assessments. The numbers may not vary greatly in a class with 20-30 students. Still, they can be crucial for large-scale exams like the one conducted by ÖSYM or the YDS exam, for instance. Maybe we can just consider calculating the average scores if we are talking about classroom assessment. Measurement errors do not tend to be very high; these are my specialty. For very small groups, things like standard deviation might not be of significant importance, in my opinion.*

TE3 expands on the relevance of statistics in assessment for classroom teachers to emphasize the importance of understanding what the average score represents and how this understanding can be used to improve the assessment process:

*There is no need for it. There are many programs that calculate statistical numbers. But when does it become significant? For example, concepts like standard deviation and measurement error might be considered when taking the average score of all senior-year high school students in English lessons. However, for a class with 20-30 students, they do not need to know much about these. Because machines and programs calculate them for us anyway. [...] The teacher needs to only know how to interpret these numbers. Interpretation, like what does that number represent? The class average is 35. What does that mean? Maybe you have not taught well enough or done a wrong assessment.*

*There could be errors in your assessment tools. The teacher can use this information to improve the assessment processes.*

When the interview excerpts are examined in light of the quantitative data findings, it can be concluded that the items allocated to the *Technical Skills* dimension do not accurately represent TEs' conceptualization of this dimension. While most of the items reserved for the *Technical Skills* dimension were considered important by the TEs, there is a possibility that their understanding of these items was different from anticipated. For instance, item 15, "having knowledge of how language testing can be used appropriately in the classroom," was associated with the practical aspects of assessment rather than technical issues by the TEs, as qualitative findings suggest. Similarly, knowing how to create good test items was categorized into the *Assessment-specific Knowledge* theme as well. The only item that directly corresponds to the technical issues from TEs' perspectives is item 15c, which received the lowest mean and median scores ( $M = 4.80$ ,  $Md = 4.00$ ) according to the quantitative data analysis. This issue is also evident in the scores awarded by TE1 and TE3 to item 15c, conforming to their discussion in the interviews.

### ***Pedagogical Knowledge***

*Pedagogical Knowledge* is another theme that has been directly derived from Taylor's (2013) existing dimensions. However, similar to the *Technical Skills* dimension, this theme differs from Bøhn & Tsagari' (2021) operationalization in that it incorporates the knowledge associated with formative assessment and decision-making procedures for improving the quality of teaching and learning. The washback effect was also categorized under this theme, as Bøhn & Tsagari suggested; however, according to the TEs' discussion, it has been moved under the *Assessment Principles* theme. Moreover, the distinction of issues discussed within the *Language Pedagogy* theme and the washback effect stood out from TEs' statements. Assessment in language pedagogy was generally linked to process-oriented assessment, peer and self-assessment, the role of feedback, and understanding the way assessments can be used to achieve specific learning or teaching goals. The washback effect, on the other hand, was related to the broader impact of assessments on the teaching and learning process, including how assessments can impact the way teachers teach and students learn in the classroom.

TE4 commented on this matter by highlighting the dichotomy between formative and

summative assessment, the former of which deals with awarding students grades as an end result, whereas the latter is perceived as an idealistic way of tracking the learner's development throughout the learning process:

*Our main aim here is not to use summative assessment. If we use formative assessment, which we do not as far as I observe, alternative assessment tools, such as portfolios, have become quite widespread today. This shows that the concept of formative assessment is very important. We should not actually focus on assigning grades to the student but rather tracking their development in a process individually, judging the difference between the point where they first began and the point where they have reached. [...] Therefore, if the teachers know that each student is different, each student's situation is different, and that with alternative assessment tools, they can observe the development of each student individually because, in my opinion, the most important thing in the measurement process is to give feedback to the student. If this feedback process is positive, the student will have learned something from it. Otherwise, when the student does not get feedback, we are only grading them, which is against the philosophy of student assessment.*

However, TE2 took a different stance and stressed the importance of integrating assessment approaches as opposed to fueling these competing paradigms:

*Indeed, we cannot definitively say that one type of assessment, whether it is formative or summative, is superior to the other. I believe that these should be considered as a whole. In other words, we cannot say that evaluating students in one particular way is more important. It is crucial to use different techniques, provide students with a variety of assessments, and gather as much data as possible. Then, based on students' strengths and weaknesses and their interests, we can tailor the learning process accordingly. In my view, all of these assessment methods are essential.*

TE4 and TE2's conflicting opinions on this particular issue mirror the dispute around the competing paradigms that are currently present in the educational community. TE2 further expressed concerns by supporting her discussion with the vision report MoNE published in 2022:

*In fact, the latest report published by the Ministry of National Education (MoNE) in their 2023 vision emphasizes the need for using different techniques together to reveal students' knowledge and performance. That is why I keep highlighting that if we can include not only the results but also the process, we will collect more meaningful data and respond to the students' needs more effectively. I believe that in this way, we can improve both student performance and learning processes.*

*However, we need to think very carefully about whether teachers perceive assessment as*

*solely a test or as a process. Because if we focus solely on a single test to evaluate students, assessing their performance in just one test may not be entirely suitable. That is why I believe that assessing students should involve a combination of different methods, not just one or two exams, but rather a continuous process where students are assessed over an extended period.*

In the same vein, TE5 described monitoring students' learning process as "crucial." He also expressed the need for accommodating individual differences in assessment in order to facilitate learning:

*What is crucial here is tracking where a student started and how far they have come throughout this process, being able to monitor individual differences. Fortunately, there are some excellent new techniques and studies in this field. I believe it is essential for teachers to be aware of these and be able to apply them effectively.*

Similar views were echoed by TE1, who emphasized that for assessment procedures to be successful, individual differences must be taken into account:

*The products given by the student during the process are a portfolio. I need to know which methods I can use to examine the student's development in the process, not only with the exam I do at the end of the year. There are not many different options, but at least there should be diversity because students have many individual differences. Some students like to work alone. Some students work very well in groups. When you give a group assignment to a student who likes to work alone, the student is not very successful.*

TE2 added a growing awareness of how learners' individual differences are important for tailoring assessments to their unique needs. Particularly, the remarks about the suggestions "keep e-portfolios" and "monitor students individually" imply a shift toward more personalized assessment strategies. Furthermore, her emphasis on giving adequate support to students underscores teachers' responsibility to foster student growth and development. However, the mention of the predominantly exam-oriented educational system in Turkey implies that there is still room to fully integrate these innovative assessment practices into our context:

*There are also suggestions to consider students' individual differences and even from a young age, keep e-portfolios, monitor them individually, and track their process. Accordingly, provide the necessary support to the student. It seems that efforts are being made to adapt this a bit more in the new system, but currently, we still have a predominantly exam-oriented one.*

In addition, TE2 associated the way the results and interpretations of assessment inform

teaching and learning by comparing it to a “road map” of the learning process:

*I believe that assessment is entirely integrated with learning. I think of them as a roadmap that shows us where we are in our learning journey and what we need to do. The more a teacher can integrate this process and gather data from it to adjust the learning process, the more I believe students' success will increase.*

TE5, in a similar stance, defined assessment as a “bridge” between students' current state and their potential for achievement:

*I think assessment serves as a bridge for learning. It helps us understand where and how a student can improve and succeed.*

TE1 addressed this issue by incorporating effective feedback strategies into his discussion as follows:

*We need to motivate the student, and while we give negative feedback, we need to point out the negativity, but it needs to be an encouragement. Okay, you have this potential strength, but you have to improve yourself in this area. Otherwise, if there is no encouraging attitude, the feedback you give does not carry any significance, and it does not resonate with the student. Just giving points causes the student to say okay, I received a grade, I'm not doing well, which means I'm weak; I don't need to express myself or study more.*

Thus, motivating students was another emerging topic of discussion under this theme.

TE1 expands on it as follows:

*Assessment is important, but the objective has to be to motivate the students and to show them how they can do better as well. This is why assessment is important. It occupies a critical position since we need to submit official documentation, and it is essential in terms of tracking the academic performance of the student. But our main goal has to be the improvement of the learner, motivating them, and pushing them forward to a higher level. After ensuring that assessment and evaluation are a means, not the main purpose, we need to be aware of that.*

In conclusion, the *Pedagogical Knowledge* theme covered a broad selection of assessment components while bringing attention to current debates between the paradigms of formative and summative assessment. TEs generally emphasized the role assessment plays in informing the learning process. While some informants advocated integrated and individualized assessment methods and facilitating a balanced approach to assessment, there were differing opinions regarding the importance attributed to this issue.

### ***Disciplinary Knowledge***

The *Disciplinary Knowledge* theme centers around the theoretical elements surrounding language learning, subject matter expertise, and language competences. TEs, in general, emphasize how understanding language learning theories affects teachers' LAL since they consider assessment theories to be a reflection of language learning theories. An excerpt from TE4's interview that supports this is given below:

*It is important for teachers to know language learning theories because these theories are also present in the philosophy of assessment and evaluation. If a teacher understands the constructivist approach, for example, they can apply it in the tests they create.*

He expressed how interactionism, constructivism, and sociocultural theories have pivotal roles in shaping language assessment strategies, as in the following quote:

*In fact, when we examine the framework of contemporary language learning theories, such as sociocultural theory, constructivism, interactionist, and so on, we can understand that the fundamental principle of these theories is to track the development of the learners in the process and to support the students when required. The underlying rationale behind these theories is actually to monitor the student's development in the process and to support them when needed, which is the key feature of the sociocultural and interactionist theories that we refer to as the mediator or mediation of the teacher; the guidance process is defined as supporting the student when they get stuck, which we call the zone of proximal development.*

Another salient feature introduced at this stage was the knowledge of The Common European Framework (CEFR). The quotation that follows recounts the way TE3 described his preparation and implementation of the classroom assessment course he currently teaches:

*Now, I had the students prepare questions for our classroom assessment course. What I asked them for, for example, is: "For which proficiency level did you prepare the question?" A1-A2-B1-B2. The age group is not that important, but I also pay attention to it. Knowing these concepts changes the questions you'll prepare. As you can see, you check at which level, probably B1-B2 in this case, they have prepared this question. If they knew the CEFR level, they would have to modify and simplify the text accordingly to match their level. That is why I emphasize it.*

The CEFR offers a framework of language proficiency levels to ensure that the preparation and adaptation of assessment materials align with learners' abilities and

learning outcomes. Therefore, in order to customize and determine the appropriateness of assessment tasks for the students' levels, classroom teachers are supposed to be informed about the CEFR.

*You know, there is a debate in the educational community about the criteria we apply for evaluation, whether we will adopt the lingua franca or take world English as a benchmark to determine fluency in English. Which criteria are we going to use?*

TE1 reflected on the evolving nature of English language education in a global context. Considering that English serves various purposes beyond academics, this requires educators to reconsider how they evaluate English proficiency, moving away from traditional native-like standards toward recognizing global English variations.

### ***Social Context***

The *Social Context* theme was operationalized as an overarching theme that encompassed both the *Local practices* and *Sociocultural values* dimension of Taylor's model as well as the education system of a country, including institutional demands impacting assessment practices in their particular settings. TE4, for instance, spoke of the constraints teachers face due to institutional requirements and expectations, as the following quote depicts:

*I also feel that there are some obstacles within the national education system in the same way. Even within the university systems, the same obstacles exist. We informally declare, 'There will be no final exams, but I'll be collecting assignments,' or for instance, I say, 'No midterm tests, but you'll do a presentation instead,' and then they say, 'No sir, it is mandatory to administer midterm exams.' The system always forces us to do these things. For example, you have to submit grades on certain deadlines. You have to administer a midterm exam. You have to submit grades on this date, this time, this grade, this grade, this time. I think that there is a systemic cumbersome organization.*

TE2 referred to her encounters with teachers working in public primary and secondary schools to explain how national policies affect their assessment practices.

*Especially for teachers working in the MoNE, let me also state this: During my studies, I encountered some teachers. End-of-year exam evaluations, such as university entrance exams or high school transition exams, are highly emphasized in some schools. Therefore, it is expected that teachers in these schools prepare their students more intensively for these exams. The expectations of the school's administration and the provincial directorate of national education play a crucial role. Teacher evaluations are*



*not solely based on the teacher's performance but are also influenced by the expectations of school administrators and the provincial directorate of national education, in my opinion.*

The parameters in the national education system, such as the imposition of certain types of assessment procedures, were another issue addressed in TE2's discussion.

*Actually, a significant portion of assessment in our country relies on multiple-choice tests. While there has been an increasing emphasis on performance-based assessments recently, the truth is that tests still dominate much of our assessment practices. If our tests are not reliable, our evaluation of students and their outcomes will also be unsuccessful.*

All the comments included in TEs' arguments indicate the responsibility of teachers to navigate the assessment environment in which they are teaching.

On the issue of sociocultural values, the impact of society's beliefs towards assessment is discussed in the following quotation by TE3:

*Teachers are not perceived as successful when they adopt a process-oriented system. This takes us back to the washback effect. When teachers are told, "What is the point if they can write or speak, they could not score full marks on the test." We can clearly see that the consequences of the country's national assessment system are reflected in the classroom.*

On designing assessment tasks, classroom teachers received criticism from TE3 for taking the least complicated approach available. This comment is also intricately rooted in the larger framework of institutional demands; thus, its implications should be carefully considered from both perspectives.

*TE3: "When the preparation of assessment tasks becomes a hassle, they prepare a matching task, they add one or two multiple choice questions, and the job is done. So, they choose to take the easy way out. The case is not that they do not have adequate theoretical knowledge, but from what I observed, generally more along the lines of "Who could deal with this much work?" They say that they have five classes, each consisting of 30 students, so a total of 150 students per teacher. Their excuses are like, "How will I prepare and evaluate writing exams? I do not have enough time and such."*

*Researcher: "In this case, do you think the school administration, school culture, or parents also have an impact?"*

*TE3: "Maybe, if a teacher has 150 students and they are asked to grade their exams in a week and submit the grades into the system, of course, they do."*

TE2 emphasized the importance of being informed about recent reports released by the ministry or higher education council:

*What does the MoNE or the Council of Higher Education expect from the teachers? Teachers should be familiar with the regulations because these regulations are constantly updated based on global developments.*

### ***Personal Beliefs and Attitudes***

The *Personal Beliefs and Attitudes* theme takes account of the experiences and preconceptions of both teachers and students. TE2 offers insights into why teachers should be aware of these issues:

*Presenting different practical techniques to teachers and addressing certain beliefs or thoughts, such as “these will not work in our classes,” is important. However, there are certainly various process and product assessment techniques that can be used with different student groups, from the smallest to the largest classes. I believe there are techniques suitable for our students as well.*

According to TE2, while instructors may be apprehensive about the application of certain assessment methods in their own classrooms due to personal opinions and experiences, they should be encouraged to explore novel methods. Classrooms are dynamic environments that are constantly changing and evolving. If teachers let their personal beliefs become a barrier, they will be continuously challenged to be resilient and adapt to change.

*But what they observe, the experiences they have had before, and the education they have received, especially in relation to formative assessment and self-assessment, if their teachers have not included these types of assessments, students are generally unaware of these issues. Students are not familiar with both formative and self-assessment, and even peer assessment, because as far as I have noticed, teachers have always given grades to students, and when it comes to grading and assessment, unfortunately, the first thing that comes to teachers’ minds is the summative assessment philosophy.*

Effective assessment can be hindered by personal preconceptions held by both teachers and students in the classroom. Therefore, teachers should also be mindful of how students’ prior experiences shape their understanding of assessment as their familiarity with assessment approaches, such as formative assessment, self-assessment, and peer assessment, is often influenced by their previous education. Offering learners guidance where and when needed, therefore, is the teachers’ responsibility. TE4 recounted his

experience regarding as follows:

*I have some concerns about peer assessment; from time to time, when I experiment with students in class, I sometimes doubt whether they are capable of doing that evaluation. Maybe it would be more reasonable for students to receive additional training on peer assessment and then use this method, but when I ask students to perform an immediate assessment, for example, in writing classes, students usually get caught up in the technical details. Therefore, they cannot make a general evaluation and provide much help to their friends at this point. This leads to a rather static peer assessment process.*

Teachers' preferences and perceptions can also influence their assessment practices. TE4 mentioned the time and effort required for process-oriented assessments, such as scoring written assignments and giving feedback. He suggested that some teachers may prefer less time-consuming methods due to their convenience.

*Process-oriented assessments are always the most demanding issues for a teacher. It takes me about a week, maybe ten days, to mark and evaluate a single academic writing course assignment, but you know that the time and effort we put in outside the classroom, both in MoNE and elsewhere, have no financial compensation. In fact, these are things that teachers do for their own job satisfaction in terms of positioning themselves. So maybe there is a problem here, too. I mean, we should be evaluating all in-class assessments, even at the university. This should be done during class hours but not outside the class. [...] A teacher's only job is not only coming to class and lecturing. Our job involves assessment, giving feedback, students responding to this feedback, and writing and editing essays at home. In fact, it also includes extracurricular labor.*

TE3 echoed the reluctance to engage in speaking or writing assessments, which require extra effort on behalf of the teachers. He mentioned that teachers may be daunted because of the workload associated with preparing and scoring rubrics:

*But then, it is easier for them to use a multiple-choice test and then spend about half an hour evaluating it. No one wants to deal with a speaking or a writing exam because preparing a rubric is an errand, and they will have to prepare rubrics separately for both speaking and writing tasks. They will have to spend maybe five to ten minutes grading each student, so they do not bother with this type of time-consuming evaluation.*

TE1 expresses his frustration about the implementation of multiple-choice tests for the sake of practicality in the passage below:

*If the teacher administers a multiple-choice exam just for the sake of practicality and to avoid spending a lot of time grading papers, this does not help the students improve. I also could do this easily. We, too, have a huge number of students and exams to be*

*evaluated. The machines automatically scan the papers and generate the scores. It takes me a lot of time to grade the students, but if it has to be done well, I think it needs to be done appropriately. I am very much against multiple-choice tests; if at all possible, no one should use them for language assessment.*

### **3.2. The Relevancy of Taylor’s (2013) LAL Model for Turkish Teacher Educators**

The second research question aimed to determine the extent to which teacher educators found Taylor’s (2013) LAL model pertinent for EFL teachers in Türkiye. The data collected from closed-response items during the second phase were analyzed to address this question. According to the analysis, 103 of 140 responses were rated five or six on the six-point Likert scale. Therefore, it indicates that most items received high scores from the participants and were rated as either very important or important. Table 6 displays the frequency of each score assigned by the respondents.

**Table 6**

*Six-point Likert scale Scoring System*

<b>Scores</b>	<b>Definitions</b>	<b>Frequency</b>
1.00	Not at all	-
2.00	Low Importance	-
3.00	Slightly Important	6
4.00	Somewhat Important	31
5.00	Important	28
6.00	Very Important	75

As can be seen in Table 6, the majority of the items were rated very important, important, or somewhat important by the TEs in this study. None of the items were deemed as unimportant or of low importance, which indicates that the TEs in this study held the view that most of the issues related to assessment are of importance. The items with the highest scores are presented in Table 7.

**Table 7***The Highest Ranked Items from Closed-ended Questions*

<b>Dimension</b>	<b>Item</b>	<b>TE1</b>	<b>TE2</b>	<b>TE3</b>	<b>TE4</b>	<b>TE5</b>	<b>M</b>	<b>Mdn</b>
Language Pedagogy	Q13-It is important to have knowledge of self-assessment	6.00	6.00	6.00	6.00	6.00	6.00	6.00
Knowledge of Theory	Q22b-It is important to have knowledge of theories of communicative competence	5.00	6.00	5.00	6.00	6.00	5.60	6.00
Language Pedagogy	Q21-It is important to have knowledge of how assessment can be used to motivate students	5.00	6.00	5.00	6.00	6.00	5.60	6.00
Language Pedagogy	Q18-It is important to have knowledge of how assessment can be used as a diagnostic tool	4.00	6.00	6.00	6.00	6.00	5.60	6.00
Knowledge of Theory	Q17-It is important to have knowledge of different types of assessment	5.00	6.00	5.00	6.00	6.00	5.60	6.00
Language Pedagogy	Q16a-It is important to have knowledge of formative assessment	5.00	6.00	5.00	6.00	6.00	5.60	6.00
Technical Skills	Q15b-It is important to have knowledge of how good items are created	5.00	6.00	5.00	6.00	6.00	5.60	6.00

Table 7 shows that the TEs rated items 13, 21, 18, and 16a, which pertain to the *Language Pedagogy* dimension, as the highest. This outcome aligns with Taylor (2013), as it indicates that teacher educators perceive elements related to the *Language Pedagogy* dimension as particularly important. Furthermore, TEs considered an understanding of theories related to communicative competence and how good items are created as of significant importance, which differs from the model Taylor proposed for classroom teachers and aligns more closely with the framework intended for test writers and professional language testers.

In total, TEs rated the questionnaire items with a minimum score of 3, which accounted for 6 out of 140 responses. An overview of the items that the TEs assigned lower scores are presented in Table 8.

**Table 8***The Lowest Ranked Items from Closed-ended Questions*

<b>Dimension</b>	<b>Item</b>	<b>TE1</b>	<b>TE2</b>	<b>TE3</b>	<b>TE4</b>	<b>TE5</b>	<b>M</b>	<b>Mdn</b>
Sociocultural Values	Q31-It is important to have knowledge of how values in society may affect assessment	4.00	6.00	3.00	6.00	5.00	4.80	5.00
Language Pedagogy	Q30-It is important to have knowledge of how to communicate assessment results in appropriate ways (e.g., how to explain to parents the results from national tests)	4.00	6.00	4.00	6.00	4.00	4.80	4.00
Local Practices	Q26-It is important to have knowledge of local issues that may be relevant for assessment (e.g., the local educational authorities' assessment criteria, how teachers in your local district assess student performance, etc.)	4.00	6.00	4.00	6.00	4.00	4.80	4.00
Technical Skills	Q15c-It is important to have knowledge of statistical measures in order to interpret results from multiple-choice tests (e.g., mean, standard deviation, measurement error)	4.00	6.00	4.00	4.00	6.00	4.80	4.00
Local Practices	Q25-It is important to have knowledge of the Regulations to the Education Act	3.00	5.00	3.00	6.00	6.00	4.60	5.00
Knowledge of Theory	Q33-It is important to have knowledge of the history of language assessment	3.00	4.00	4.00	4.00	5.00	4.00	4.00

As can be seen from Table 8, item 31 in the *Sociocultural Values* dimension and item 30 in the *Language Pedagogy* dimension revealed differing opinions among informants, with scores ranging from 3 to 6. Similarly, item 26 in the *Local Practices* dimension and 15c in the *Technical Skills* dimension received scores between 4 and 6. Items that received somewhat lower than average scores were 25 in the *Local Practices* dimension and 33, ranging from 3 to 6. The statistics demonstrate that the TEs did not all view the aforementioned characteristics of LAL as highly important, reflecting differences in their perceived relevance.

Table 9 displays the teacher educators' responses to items associated with the *Knowledge of Theory* dimension.

**Table 9***Knowledge of Theory Dimension*

<b>Knowledge of Theory</b>	<b>TE1</b>	<b>TE2</b>	<b>TE3</b>	<b>TE4</b>	<b>TE5</b>	<b>Σ</b>	<b>M</b>	<b>Mdn</b>
Q22a-It is important to have knowledge of language learning theories	4.00	6.00	4.00	6.00	6.00	26	5.20	6.00
Q22b-It is important to have knowledge of theories of communicative competence	5.00	6.00	5.00	6.00	6.00	28	5.60	6.00
Q27-It is important to have knowledge of assessment/testing theory (related to «validity, » «reliability, » etc.)	4.00	6.00	6.00	4.00	6.00	26	5.20	6.00

According to Table 9, TEs tend to lean toward agreement on the significance of having knowledge of language learning theories ( $M=5.20$ ), theories of communicative competence ( $M = 5.60$ ), and assessment/testing theories ( $M = 5.20$ ) within the *Knowledge of Theory* dimension, as high mean and median (6.00) values for these items indicate.

Table 10 illustrates the TE's responses to the questions pertaining to the *Technical Skills* dimension.

**Table 10***Technical Skills Dimension*

<b>Technical Skills</b>	<b>TE1</b>	<b>TE2</b>	<b>TE3</b>	<b>TE4</b>	<b>TE5</b>	<b>Σ</b>	<b>M</b>	<b>Mdn</b>
Q15a-It is important to have knowledge of how language testing can be used appropriately in the English language	3.00	6.00	4.00	6.00	6.00	25	5.00	6.00

classroom, e.g., multiple choice								
Q15b-It is important to have knowledge of how good items are created	5.00	6.00	5.00	6.00	6.00	28	5.60	6.00
Q15c-It is important to have knowledge of statistical measures in order to interpret results from multiple-choice tests (e.g., mean, standard deviation, measurement error)	4.00	6.00	4.00	4.00	6.00	24	4.80	4.00

The overall scores within the *Technical Skills* dimension indicate that the TEs attribute slightly greater importance to technical know-how compared to what Taylor (2013) suggested for classroom teachers. Informants in the present study particularly emphasized the significance of skills related to understanding the process of creating effective test items ( $M = 5.60$ ) compared to the other aspects of this dimension.

Table 11 details the scores TEs assigned for the item related to the *Principles and Concepts* dimension.

**Table 11**

*Principles and Concepts Dimension*

<b>Principles and Concepts</b>	<b>TE1</b>	<b>TE2</b>	<b>TE3</b>	<b>TE4</b>	<b>TE5</b>	<b>Σ</b>	<b>M</b>	<b>Mdn</b>
Q28-It is important to have knowledge of ethical issues (fairness, use of assessment results for purposes other than what was intended, etc.)	4.00	6.00	5.00	6.00	6.00	27	5.40	6.00



Table 11 shows that the scores for item 28 range from 4 to 6, with a mean of 5.40 and a median of 6.00. These scores reflect a shared view among the respondents, who place high importance on ethical issues, as Taylor (2013) anticipated.

Table 12 provides an overview of TEs' responses to items within the *Language Pedagogy* dimension.

**Table 12**  
*Language Pedagogy Dimension*

<b>Language Pedagogy</b>	<b>TE1</b>	<b>TE2</b>	<b>TE3</b>	<b>TE4</b>	<b>TE5</b>	<b>Σ</b>	<b>M</b>	<b>Mdn</b>
Q12-It is important to have knowledge of how assessment can promote learning	4.00	6.00	5.00	6.00	6.00	27	5.40	6.00
Q13-It is important to have knowledge of self-assessment	6.00	6.00	6.00	6.00	6.00	30	6.00	6.00
Q14-It is important to have knowledge of peer-assessment	4.00	6.00	6.00	5.00	6.00	27	5.40	6.00
Q16a-It is important to have knowledge of formative assessment	5.00	6.00	5.00	6.00	6.00	28	5.60	6.00
Q18-It is important to have knowledge of how assessment can be used as a diagnostic tool	4.00	6.00	6.00	6.00	6.00	28	5.60	6.00
Q21-It is important to have knowledge of how assessment can be used to motivate students	5.00	6.00	5.00	6.00	6.00	28	5.60	6.00
Q30-It is important to have knowledge of how to communicate assessment results in appropriate ways (e.g., how to explain to parents the results from national tests)	4.00	6.00	4.00	6.00	4.00	24	4.80	4.00

According to the data provided, informants uniformly emphasized the role of self-assessment (Q13), awarding it the highest possible rating of 6.00. Formative assessment (Q16a), the use of assessment as a diagnostic tool (Q18), and the use of assessment to motivate students (Q21) also emerged as key elements within this dimension. TEs gave these items scores that ranged from 5 to 6, with a mean of 5.60 and a median of 6.00, suggesting an overall agreement on their importance for LAL. However, the item regarding effective communication of assessment results (Q30) reveals varying perspectives among the TEs, with a mean score of 4.80 and a median score of 4.00, suggesting that the importance of this aspect was perceived as somewhat lower within the *Language Pedagogy* dimension.

Table 13 exhibits how TEs rated the item associated with the *Sociocultural Values* dimension.

**Table 13**

*Sociocultural Values Dimension*

<b>Sociocultural Values</b>	<b>TE1</b>	<b>TE2</b>	<b>TE3</b>	<b>TE4</b>	<b>TE5</b>	<b>Σ</b>	<b>M</b>	<b>Mdn</b>
Q31-It is important to have knowledge of how values in society may affect assessment	4.00	6.00	3.00	6.00	5.00	24	4.80	5.00

Table 13 demonstrates that informants found item 31, pertaining to Taylor's *Sociocultural Values* dimension, to be relatively less relevant than the proposed model. However, this is not an indication that they were considered unimportant by the TEs. As the table demonstrates, this component received a diversity of ratings from 3 to 6, with a mean of 4.80 and a median of 5.00.

Table 14 displays how TEs perceive the *Local Practices* dimension in Taylor's model.

**Table 14***Local Practices Dimension*

<b>Local Practices</b>	<b>TE1</b>	<b>TE2</b>	<b>TE3</b>	<b>TE4</b>	<b>TE5</b>	<b>Σ</b>	<b>M</b>	<b>Mdn</b>
Q24-It is important to have knowledge of the national curriculum (the general part and the English subject curriculum)	6.00	5.00	4.00	6.00	6.00	27	5.40	6.00
Q25-It is important to have knowledge of the Regulations to the Education Act	3.00	5.00	3.00	6.00	6.00	23	4.60	5.00
Q26-It is important to have knowledge of local issues that may be relevant for assessment (e.g., the local educational authorities' assessment criteria, how teachers in your local district assess student performance, etc.)	4.00	6.00	4.00	6.00	4.00	24	4.80	4.00

While the item related to the national curriculum is regarded as relatively more important than the other elements, such as knowledge of education regulations and awareness of local assessment criteria, had lower levels of importance attributed to them. Particularly, item 24 was considered to be the most important aspect within the *Local Practices* dimension. The scores for this item varied from 4 to 6, resulting in a mean score of 5.40 and a median of 6.00. These scores indicate the significance of understanding the national curriculum, particularly the English subject curriculum.

Table 15 shows the ratings for the *Personal Beliefs and Attitudes* dimension.

**Table 15***Personal Beliefs and Attitudes Dimension*

<b>Personal Beliefs and Attitudes</b>	<b>TE1</b>	<b>TE2</b>	<b>TE3</b>	<b>TE4</b>	<b>TE5</b>	<b>Σ</b>	<b>M</b>	<b>Mdn</b>
Q32-It is important to have knowledge of how one's own view on assessment (values, outlook) may affect assessment	4.00	6.00	3.00	6.00	6.00	25	5.00	6.00

The TEs in this study gave item 32 scores ranging from 3 to 6, with a mean score of 5.00 and a median of 6.00. The result described above is consistent with Taylor's framework for classroom teachers.

Table 16 provides an outline of the *Scores and Decision-Making* dimension within Taylor's model.

**Table 16***Scores and Decision-Making Dimension*

<b>Scores and Decision-Making</b>	<b>TE1</b>	<b>TE2</b>	<b>TE3</b>	<b>TE4</b>	<b>TE5</b>	<b>Σ</b>	<b>M</b>	<b>Mdn</b>
Q16b-It is important to have knowledge of summative assessment	5.00	6.00	4.00	5.00	6.00	26	5.20	5.00
Q29-It is important to have knowledge of how to use rating scales/scoring rubrics	4.00	6.00	5.00	6.00	6.00	27	5.40	6.00

The scores for the components in this dimension, along with those in the *Knowledge of Theory* and *Principles and Concepts* dimensions, differed considerably from what Taylor (2013) proposed for classroom teachers' LAL profile. As can be seen in Table R10, item

16b received ratings from 4 to 6, with a mean score of 5.20 and a median of 5.00, and item 29 ranged from 4 to 6, with an overall mean score of 5.40 and a median of 6.00. The relatively high mean and median scores of these items indicate that the TEs place high importance on the *Scores and Decision-Making* dimension, which extends to both summative assessment and the usage of rating scales/scoring rubrics.

In order to make the results comparable to Taylor's (2013) and Bøhn and Tzagari's (2021) models, a statistical software program was used to compute the average median scores of the items in each dimension. The values Taylor (2013) originally established for teachers' LAL profiles and the conversion table used for this study are presented in Table 17 and Table 18, respectively.

**Table 17**

*Taylor's (2013) LAL Levels for Classroom Teachers*

<b>Dimensions</b>	<b>Values</b>
Knowledge of Theory	2
Technical Skills	3
Principles and Concepts	2
Language Pedagogy	4
Sociocultural Values	3
Local Practices	3
Personal Beliefs and Attitudes	3
Scores and Decision-Making	2

**Table 18**

*Conversion of Average Median Scores*

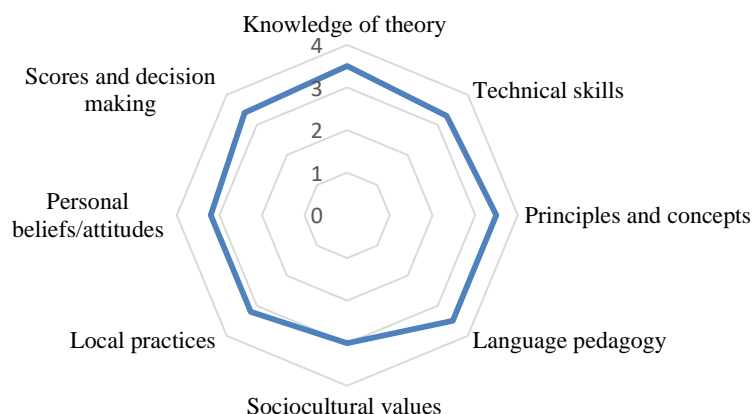
<b>Dimensions</b>	<b>Average Median Values</b>	<b>Converted Median Values</b>
Knowledge of Theory	5.40	3.52
Technical Skills	5.20	3.36
Principles and Concepts	5.40	3.52
Language Pedagogy	5.40	3.52
Sociocultural Values	4.80	3.04
Local Practices	5.00	3.20
Personal Beliefs and Attitudes	5.00	3.20

As Bøhn and Tsagari (2021) previously reported, the values in Taylor's LAL model for the stakeholders were established on a five-point scale, while the questionnaire items in both the present study and Bøhn and Tsagari's study were measured using a six-point Likert scale. Therefore, with the aim of equalizing the results, the same conversion formula Bøhn and Tsagari used ( $0.8x - 0.8$ ) was applied to the average median scores.

In light of these results, Figure 9 illustrates the LAL profile for classroom teachers based on TEs' perceptions in the Turkish context.

**Figure 9**

*Classroom Teachers' LAL Profile from Turkish Teacher Educators' Perspective*



As shown in Figure 9, the average median values in the present study tended to align closely with Taylor's hypothesized levels for *Language Pedagogy*, *Local Practices*, *Personal Beliefs and Attitudes*, and *Sociocultural Values* dimensions. However, notable variations can be observed in the following dimensions: *Knowledge of Theory*, *Technical Skills*, *Principles and Concepts*, and *Scores and Decision-Making*.

## 4. DISCUSSION AND CONCLUSION

In this section, conclusions drawn from the study findings are presented and discussed in line with the related studies. The current investigation focused on two facets of teacher educators' perceptions of LAL: First, the conceptualization of LAL involved exploring how Turkish teacher educators perceived LAL. Then, the relevance of Taylor's (2013) LAL model was analyzed to reveal its applicability within the Turkish context.

### 4.1. Classroom Teachers' LAL Profile from Teacher Educators' Perspectives

In their study on Norwegian teacher educators' conceptions of LAL, Bøhn and Tsagari (2021) identified a limitation in Taylor's (2013) model, noting that it lacked detailed descriptions of LAL dimensions. The present research, therefore, had its basis in the call for the empirical validation of this model and its relevance within a different context. The qualitative findings obtained from the first research question yielded the following dimensions of classroom teachers' LAL profile:

- Theoretical Knowledge of Assessment
- Competence in Scoring and Statistics
- Language Pedagogy Knowledge
- Content Knowledge
- Social Context Knowledge

#### *Theoretical Knowledge of Assessment*

The first dimension, Theoretical Knowledge of Assessment, which was identified in line with the findings of the present study, encompasses the concepts specified in Taylor's (2013) model under the *Knowledge of Theory and Principles and Concepts* dimensions. Additionally, this category also aligns with a broader one identified as the Knowledge of Assessment Theory dimension in Bøhn and Tsagari's (2021) study. Both of these dimensions involve assessment theories and concepts, such as validity, reliability, and washback. However, this dimension has a greater focus on understanding the purposes of assessment and the significance of selecting appropriate assessment methods and instruments suitable for them. Furthermore, it addresses elements of Brindley's (2001)

professional development model, particularly those related to the *Defining and Describing Proficiency* module. These elements involve theoretical foundations of language testing and assessment procedures, such as how to measure communicative competence and how to develop assessment tools in line with the communicative approaches to second language teaching methodologies. The aforementioned issues resonate closely with Inbar-Lourie's (2008, p. 390) LAL framework, namely the reason for conducting an assessment, the characteristics to be assessed, and the form of the assessment.

The TEs argued that when teachers lack theoretical knowledge about assessment, their evaluations are likely to be unreliable or inadequate. First and foremost, they agreed on the critical importance of understanding assessment philosophy for the reason that they regarded this to be the key to becoming assessment literate and being conscious of the decisions made while delivering assessments. Likewise, participants emphasized the necessity for teachers to be knowledgeable about the practical aspects of assessment, for instance, selecting appropriate assessment methods and strategies for their particular contexts and course objectives and designing or modifying assessment tasks that align with the aims and objectives that have been identified.

Another area addressed by the TEs concerned assessment principles and concepts. TEs mentioned that teachers should be well aware of concepts such as validity, reliability, and authenticity. Their opinions on these matters were closely related to the components of test usefulness elucidated by Bachman and Palmer (1996). Specifically, TEs' discussion touched upon authenticity, referring to teachers' competence in ensuring properties of the assessment tasks reflected those of real-world tasks and interactiveness as teachers' awareness of how individuals' traits, including language ability, background knowledge, and motivation, influenced the assessment procedures.

In addition to gaining these core competencies, TEs also stressed the necessity of adopting a lifelong learning approach to remain informed and literate in the language assessment field (Inbar-Lourie, 2008). According to the respondents, teachers need to be constantly up-to-date with the latest developments in language assessment, theories of assessment and language learning, technological innovations, and relevant research in this domain. In conclusion, the TEs' perspectives confirmed that teachers, particularly those in



the pre-service education stage, need to gain a solid foundation in assessment theories, current techniques, and alternative resources in order to carry out assessments promoting effective language learning.

### ***Competence in Scoring and Statistics***

Competence in Scoring and Statistics dimension primarily focuses on the implications of the *Technical Skills* dimension from Taylor's (2013) LAL model. The findings from the qualitative analysis suggested that some items associated with *Technical Skills*, such as knowledge of how language testing can be used appropriately in the classroom and creating good test items, were regarded as practical aspects of assessment and were associated with the components of the 'Theoretical Knowledge of Assessment' category mentioned above. Informants in the present study argued that classroom teachers rarely require complex statistical knowledge in their routine assessment practices. TEs generally believed that understanding practical aspects, such as choosing among various assessment purposes and preparing assessment tasks based on existing criteria, was more important than acquiring statistical expertise. While this perspective aligned with Brindley's (2001) argument in favor of offering a criterion-referenced training module for classroom teachers rather than a statistical one, the contrasting remarks made by TE2 echoed those of Inbar-Lourie (2008), who also strongly emphasized that statistical elements cannot be neglected and that classroom teachers need to be well-informed about test development and analysis. The analysis of overall responses in the interviews revealed that TEs thought classroom teachers should refrain from using assessment or test tasks that require statistical skills, such as multiple-choice items, in their routine classroom-based assessments. They believe that assessment measures such as multiple-choice tests are inadequate in eliciting a true reflection of learners' language abilities. These insights clearly demonstrate the major steps taken to implement formative assessment practices in Türkiye.

At this stage, we came up with the suggestion to examine the components related to Taylor's (2013) *Scores and Decision-Making* dimension under two separate categories as 'decision-making' and 'scoring.' Although TEs in this study argued that classroom teachers may not need to engage in complex statistical calculations, they still highlighted the importance of being competent in interpreting the results of assessment data. They also

suggested that in order for teachers to form sound decisions about their teaching strategies and pinpoint areas where students' performances need improvement, they must understand what basic statistics analysis and interpretations, such as score averages, indicate. These findings imply that TEs in Türkiye interpreted the items within the *Scores and Decision-Making* dimension as making formative judgments rather than the statistical aspects of scoring procedures. Consequently, the decision-making aspect of Taylor's (2013) *Scores and Decision-Making* dimension was associated with teachers' language pedagogy knowledge.

Bøhn and Tsagari (2021) explored the relevance of Taylor's (2013) LAL profile for classroom teachers from the perspectives of Norwegian teacher educators. Upon reviewing their investigation, similar outcomes were found. They suggested that Taylor's (2013) *Scores and Decision-Making* dimension should be renamed as 'Scoring' and aspects related to decision-making should be moved to the *Language Pedagogy* dimension. However, they maintained that the *Technical Skills* dimension required no revisions in contrast to our argument to incorporate its components under the 'Theoretical Knowledge of Assessment' category.

### ***Language Pedagogy Knowledge***

Language Pedagogy Knowledge, therefore, relates to the role of assessment in aiding learners' development and improving teaching and learning. TEs in the present study stated that assessment should not be the primary focus but rather a way for teachers to facilitate learning. This perspective is evident from their descriptions of assessment as a *roadmap* or a *bridge* between learners' current levels and potential. Moreover, TEs argued that the data collected through assessment results should be utilized as a valuable source that informs teaching practices to meet students' needs. They also recommended adopting an integrated approach to language assessment and encouraging teachers to recognize the advantages of both formative and summative assessment. In the same vein, the informants emphasized the significance of addressing learners' individual differences in assessment practices. They underscored the importance of adapting assessment methods to accommodate various learning styles and preferences and using personalized assessment methods to focus on learners' strengths and weaknesses. Effective feedback strategies that motivate students and

provide encouragement were another issue discussed within this theme. The role of teachers in encouraging their students to attain higher levels of success was emphasized.

A review of the current study's findings and those of Bøhn and Tsagari's (2021) study suggests that this dimension does not require changes in terms of its content. However, the present study incorporated Shulman's (1987) notion of pedagogical content knowledge into the title of this dimension because, as mentioned previously in the first chapter, this term reflects the interrelationship between content knowledge and pedagogical knowledge.

### ***Content Knowledge***

A new categorization called Content Knowledge was formed based on the findings of Bøhn and Tsagari's (2021) study. This category encompasses aspects pertaining to disciplinary knowledge, including familiarity with language learning theories, approaches, and methods like second language acquisition, communicative competence, constructivist theory, CEFR, Standardized English, English as a world language, and English as a lingua franca.

Findings of the study revealed that TEs saw a direct relationship between assessment methods and their underlying language-learning theories. According to the participants, theories such as constructivism and sociocultural theory essentially shaped educational assessment practices. Thus, they suggested that a strong theoretical background in language education may guide teachers in carrying out good assessments. Additionally, knowledge of the CEFR was regarded as another essential aspect of language assessment in terms of meeting international assessment standards. Moreover, TEs found this particularly relevant for ensuring that teaching and assessment procedures are aligned with international criteria for language proficiency.

### ***Social Context Knowledge***

Finally, the Social Context Knowledge category involves institutional constraints teachers often encounter. These constraints include testing practices and expectations imposed by educational institutions and school administrators. TEs believe that these demands can limit teachers' flexibility in designing assessments that align with their students' needs. Moreover, they believe that national education policies and standardized tests have a significant impact on Turkish teachers' assessment practices. They feel that

classroom language teachers are often tasked with student preparation for student selection exams, which leads to an overemphasis on test preparation and negative washback. According to the informants, efforts to adhere to national assessment standards limit teacher autonomy. In addition, teachers frequently encounter time and resource constraints when designing and implementing assessments. Large class sizes, heavy workloads, and limited time for marking student work can lead to the use of convenient but less effective assessment methods, such as multiple-choice tests. These constraints may discourage teachers from adopting more innovative assessment approaches.

The dimension of *Sociocultural Values* was another aspect in Taylor's (2013) model that was also called into question by Bøhn and Tsagari (2021), although neither of these studies provided a detailed account of the elements this dimension entails. As a result of this, there appeared to be no explicit emphasis on the societal variables in the quantitative analysis. Although not addressed explicitly, the present investigation's qualitative results indicated that TEs discussed how values in society shaped teachers' assessment practices within each LAL dimension. For instance, while discussing teachers' personal beliefs and institutional constraints, TEs brought up sociocultural values by reflecting on how society's perspectives influence education, assessment, and language learning. In the Turkish context, it was argued that our society highly valued standardized testing and equated it with success. An example could be provided by investigating the respondents' perspectives on the globalization of English. This issue can both be an impact of sociocultural value and a contextual issue, depending on how it is articulated. If societal values about English as a global lingua franca are discussed, it may be associated with sociocultural values. However, the issue becomes a contextual concern if it pertains to the manner in which this global pattern affects assessment practices in a particular context. This is also reflected in respondents' perceptions of the washback effect. The washback effect may be discussed as a component of sociocultural values if it refers to how assessments driven by societal expectations impact teaching and learning methods. For instance, if society places a great deal of importance on academic achievement, teachers may adopt a "teaching to the test" philosophy that hinders formative assessment practices.

#### 4.2. Comparing Taylor's (2013) and Turkish Educators' Perspectives on Classroom Teachers' LAL Profiles

A comparison of Taylor's (2013) LAL profile for classroom teachers and the levels that emerged from the current investigation is presented in Figure 10.

##### Figure 10

*Comparing Taylor's (2013) and Turkish Educators' Perspectives on Classroom Teachers' LAL Profiles*

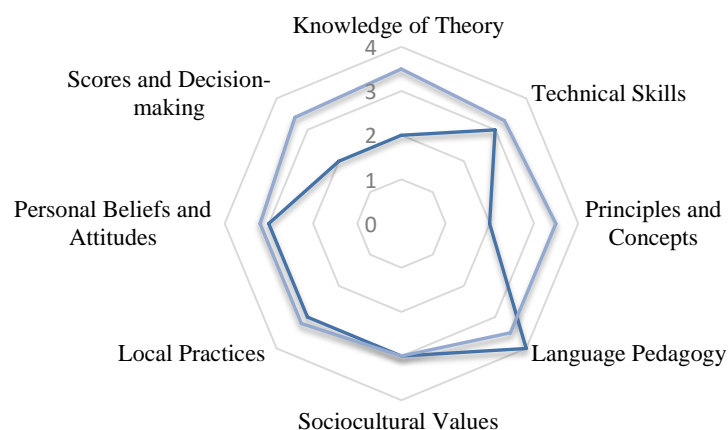


Figure 10 functions as a model for judging the extent to which participants in the present study found the LAL profile for classroom teachers relevant. As demonstrated in Figure 10, while Taylor (2013) assigned a value of 2.00 for the *Knowledge of Theory* dimension, the overall score TEs in this study gave this dimension was 3.52. This indicates that informants in this study perceived the aspects related to measurement theory to be considerably more important than Taylor predicted. In the same vein, the *Principles and Concepts* dimension received a higher value (3.52) than Taylor's (2.00), suggesting that TEs considered ethical issues to be a more relevant area for classroom teachers. Similarly, the *Scores and Decision-Making* dimension was deemed of greater significance, with a value of 3.44 as opposed to a value of 2 in Taylor (2013). The *Technical Skills* dimension was also found somewhat more important than what Taylor anticipated, with a value of 3.36. A high

degree of agreement is observed among TEs in the study and Taylor's model regarding the significance of *Language Pedagogy*, *Local Practices*, *Personal Beliefs and Attitudes*, and *Sociocultural Values* dimensions, as the overall scores for these components correspond closely to those in Taylor's model.

TEs rated items 13, 21, 18, and 16a, which pertain to the *Language Pedagogy* dimension, as the highest. This outcome aligns with both Taylor's (2013) and Bøhn and Tsagari's (2021), as it indicates that teacher educators perceive elements related to the *Language Pedagogy* dimension as particularly important. When the interview excerpts are examined in light of the quantitative data findings, it can be concluded that the items categorized under *Technical Skills* do not accurately represent TEs' conceptualization of this dimension. For instance, item 15, "having knowledge of how language testing can be used appropriately in the classroom," was associated with the practical aspects of assessment rather than technical issues by the TEs, as qualitative findings suggest. Similarly, knowing how to create good test items was categorized into the 'Theoretical Knowledge of Assessment' theme as well. The only item that directly corresponds to the technical issues based on TEs' perspectives is item 15c, which received the lowest mean and median scores ( $M = 4.80$ ,  $Mdn = 4.00$ ) according to the quantitative data analysis. This issue can be solved by separating statistical aspects of assessment from the knowledge of assessment concepts and principles. Moreover, TEs' understanding of the *Scores and Decision-making* dimension revealed that making decisions based on assessment outcomes were associated with the *Language Pedagogy* dimension. These results highlight the need for a more accurate representation of Taylor's (2013) LAL dimensions in future research.

### **Implications**

The findings of this study have several implications for the development of teacher LAL. First of all, teacher educators point to the importance of strengthening teacher training programs to foster pre-service and in-service teachers' assessment competencies in areas such as assessment theory, scoring and statistical measures, language pedagogy, and social and contextual issues in order to empower them in designing, selecting, and adapting sound assessment tasks (Inbar-Lourie, 2008; Kremmel & Harding, 2019). Findings also indicate that assessment concepts such as washback should be emphasized in terms of

social and contextual issues in teacher education and professional development programs. Moreover, they suggest that it is important for these programs to integrate various assessment methods, techniques, and strategies. Additionally, this study, conducted with teacher educators who are valuable stakeholders in the language assessment field, has demonstrated the importance of investigating different stakeholders' perspectives in conceptualizing LAL, calling attention to the urgent need for establishing a comprehensive LAL culture that meets the requirements of the language assessment community.

## REFERENCES

- Alderson, J. C. (2005). *Diagnosing foreign language proficiency: The interface between learning and assessment*. New York, NY: Continuum.
- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14(2), 115–129. <https://doi.org/10.1093/applin/14.2.115>
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford University Press.
- Bachman, L. F. (2000). Modern language testing at the turn of the century: assuring that what we count counts. *Language Testing*, 17(1), 1–42. <https://doi.org/10.1177/026553220001700101>
- Bachman, L.F. & Palmer, A.S. (1996) *Language Testing in Practice: Designing and Developing Useful Language Tests*. Oxford University Press, Oxford.
- Bachman, L., & Damböck, B. (2018). *Language assessment for classroom teachers*. Oxford University Press.
- Bailey, K. M. (1996). Working for washback: a review of the washback concept in language testing. *Language testing*, 13(3), 257–279.
- Bøhn, H., & Tzagari, D. (2021). Teacher educators' conceptions of language assessment literacy in Norway. *Journal of Language Teaching and Research*, 12(2), 222–233. <https://doi.org/10.17507/jltr.1202.02>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Brindley, G. (1998). Outcomes-based assessment and reporting in language learning programmes: A review of the issues. *Language Testing*, 15(1), 45–85. <https://doi.org/10.1177/026553229801500103>
- Brindley, G. (2001). Language assessment and professional development. In C. Elder, A. Brown, K. Hill, N. Iwashita, T. Lumley, T. McNamara, & K. O'Loughlin (Eds.), *Experimenting with uncertainty: Essays in honour of Alan Davies* (pp. 126–36). Cambridge, UK: Cambridge University Press.



- Brookhart, S. M., & DeVoge, J. G. (1999). Testing a theory about the role of classroom assessment in student motivation and achievement. *Applied Measurement in Education*, 12(4), 409–425. [https://doi.org/10.1207/S15324818AME1204\\_5](https://doi.org/10.1207/S15324818AME1204_5)
- Brookhart, S. M. (1999) *The art and science of classroom assessment: The missing part of the pedagogy* (ASHE-ERIC Higher Education Report, 27(1)) (Washington, DC, George Washington University).
- Brown, H. D. (2004). *Language Assessment: Principles and Classroom Practices*. New York: Longman.
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1(1), 1–47. <http://dx.doi.org/10.1093/applin/I.1.1>
- Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge University Press.
- Coniam, D. (2009). Investigating the quality of teacher-produced tests for EFL students and the effects of training in test development principles and practices on improving test quality. *System*, 37(2), 226-242. <https://doi.org/10.1016/j.system.2008.11.008>
- Crusan, D. J., Plakans, L., & Gebriel, A. (2016). Writing assessment literacy: Surveying second language teachers' knowledge, beliefs, and practices. *Assessing Writing*, 28, 43–56. <https://doi.org/10.1016/j.asw.2016.03.001>
- Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing Among Five Approaches*. SAGE.
- Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing*, 25(3), 327–347. <https://doi.org/10.1177/0265532208090156>
- European Commission, Directorate-General for Education, Youth, Sport and Culture, Siarova, H., Mašidlauskaitė, R., Sternadel, D. (2017). *Assessment practices for 21st century learning: review of evidence: analytical report*, Publications Office. <https://data.europa.eu/doi/10.2766/71491>

- Fulcher, G. and Davidson, F. (2007). *Language testing and assessment: An advanced resource book*. London: Routledge.
- Falsgraf, C. (2005). Why a national assessment summit? New visions in action. National Assessment Summit. Meeting conducted in Alexandria, Va. Retrieved from <https://files.eric.ed.gov/fulltext/ED527580.pdf>
- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113–132. <https://doi.org/10.1080/15434303.2011.642041>
- Fulcher, G., & Davidson, F. (2007). *Language testing and assessment: An advanced resource book*. Oxon: Routledge.
- Gipps, C., Brown, M., McCallum, B., & McAlister, S. (1995). *Intuition or evidence? Teachers and national assessment of seven-year-olds*. Buckingham, UK: Open University Press.
- Giraldo, F. (2020). A post-positivist and interpretive approach to researching teachers' language assessment literacy. *Profile: Issues in Teachers' Professional Development*, 22(1), 189–200. <https://doi.org/10.15446/profile.v22n1.78188>
- Harding, L., & Kremmel, B. (2016). Teacher assessment literacy and professional development. In D. Tsagari & J. Banerjee (Eds.), *Handbook of Second Language Assessment* (pp. 413–427). De Gruyter Mouton.
- Hatipoğlu, Ç. (2015). English language testing and evaluation (ELTE) training in Turkey: Expectations and needs of pre-service English language teachers. *ELT Research Journal*, 4 (2), 111–128. Retrieved from <https://dergipark.org.tr/tr/pub/eltrj/issue/28780/308006>
- Huerta-Macias, A. (1995). Alternative assessment: responses to commonly asked questions. *TESOL Journal*, 5(1), 8–11.
- Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*, 25(3), 385–402. <https://doi.org/10.1177/0265532208090158>

- Jeong, H. (2013). Defining assessment literacy: Is it different for language testers and non-language testers? *Language Testing*, 30(3), 345–362. <https://doi.org/10.1177/0265532213480334>
- Kaiser, G., & Willander, T. (2005). Development of mathematical literacy: Results of an empirical study. *Teaching Mathematics and Its Applications*, 24(2–3), 48–60. <https://doi.org/10.1093/teamat/hri016>
- Kremmel, B., & Harding, L. (2019). 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>
- Kunnan, A. J. (2004). Test fairness. In M. Milanovic, & C. Weir (Eds.), *European language testing in a global context*, (pp. 27–48). Cambridge: CUP.
- Lam, R. (2014). Language assessment training in Hong Kong: Implications for language assessment literacy. *Language Testing*, 32(2), 169–197. <https://doi.org/10.1177/0265532214554321>
- McMillan, J. H. (2000). Fundamental assessment principles for teachers and school administrators. *Practical Assessment, Research and Evaluation*, 7(8), 1–5. <https://doi.org/10.7275/5kc4-jy05>
- McMillan, J. H., & Nash, S. (2000). *Teacher classroom assessment and grading practices decision making* [Paper presentation]. 2000 Annual Meeting of the National Council on Measurement in Education, New Orleans.
- Mede, E., & Atay, D. (2017). English language teachers' assessment literacy: The Turkish context. *Dil Dergisi*, 168(1), 1–5. [https://doi.org/10.1501/Dilder\\_0000000237](https://doi.org/10.1501/Dilder_0000000237)
- Messick, S. (1989). Meaning and values in test validation: *The Science and Ethics of Assessment*. *Educational Researcher*, 18(2), 5–11. <https://doi.org/10.3102/0013189X018002005>
- Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13(3), 241–256. <https://doi.org/10.1177/026553229601300302>

Ministry of National Education [MoNE] (2022). *Milli eğitim istatistikleri: örgün eğitim* [National education statistics: formal education].

Retrieved from  
[https://sgb.meb.gov.tr/meb\\_iys\\_dosyalar/2022\\_09/15142558\\_meb\\_istatistikleri\\_orgun\\_e\\_gitim\\_2021\\_2022.p](https://sgb.meb.gov.tr/meb_iys_dosyalar/2022_09/15142558_meb_istatistikleri_orgun_e_gitim_2021_2022.p)

Moustakas, C. (1994). *Phenomenological research methods*. SAGE.

Morrow, K. (2004). *Insights from the Common European Framework*. Oxford: Oxford University Press.

OECD. (2019). *Education at a Glance 2019 OECD indicators: OECD Indicators*. OECD Publishing.

O'Loughlin K. (2006). Learning about second language assessment: Insights from a postgraduate student on-line forum. *University of Sydney Papers in TESOL*, 1(1), 71–85.

O'Loughlin, K. (2013). Developing the assessment literacy of university proficiency test users. *Language Testing*, 30(3), 363-380. <https://doi.org/10.1177/0265532213480336>

Önalın, O., & Gürsoy, E. (2020). EFL teachers' views and needs on in-service training as a part of professional development: A case study in Turkish context. *Bartın University Journal of Faculty of Education*, 9(2), 373–387. DOI: 10.14686/buefad.713100

Patton, M. (1990). *Qualitative Evaluation and Research Methods*. Beverly Hills, CA: Sage.

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/10.1177/0265532213480>

Scarino, A. (2013). Language assessment literacy as self-awareness: Understanding the role of interpretation in assessment and in teacher learning. *Language Testing*, 30(3), 309–327. <https://doi.org/10.1177/0265532213480128>

Sevimel-Şahin, A., & Subaşı, G. (2019). An overview of language assessment literacy research within English language education context. *Journal of Theoretical Educational Science*, 12 (4), 1340–1364. <https://doi.org/10.30831/akukeg.501817>

- Sevimel-Şahin, A., & Subaşı, G. (2021). Exploring foreign language assessment literacy training needs of pre-service English language teachers. *International Online Journal of Education and Teaching (IOJET)*, 8(4), 2783–2802. Retrieved from <https://iojet.org/index.php/IOJET/article/view/1478>
- Shulman, L.S. (1986) Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15, 4–14. <http://dx.doi.org/10.3102/0013189X015002004>
- Shulman, L.S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57(1), 1–23. <https://doi.org/10.17763/haer.57.1.j463w79r56455411>
- Stabler-Havener, M. L. (2018). Defining, conceptualizing, problematizing, and assessing language teacher assessment literacy. *Studies in Applied Linguistics and TESOL*, 18(1), 1–22. <https://doi.org/10.7916/salt.v18i1.1195>
- Stiggins, R. J. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan*, 83(10), 758–765. <https://doi.org/10.1177/003172170208301010>
- Taylor, L. (2009). Developing assessment literacy. *Annual Review of Applied Linguistics*, 29, 21–36. <https://doi.org/10.1017/S0267190509090035>
- Taylor, L. (2013). Communicating the theory, practice and principles of language testing to test stakeholders: Some reflections. *Language Testing*, 30(3), 403–412. <https://doi.org/10.1177/0265532213480338>
- Tsagari, D., & Vogt, K. (2017). Assessment literacy of foreign language teachers around Europe: Research, challenges and future prospects. *Papers in Language Testing and Assessment*, 6(1), 41–63. <https://doi.org/10.58379/uhix9883>
- Viengsang, R. (2016). Exploring pre-service English teachers' language assessment literacy. *Modern Journal of Language Teaching Methods*, 6(5), 432–442. Retrieved from <http://mjltm.org/article-1-63-en.html>
- Vogt, K., & Tsagari, D. (2014). Assessment literacy of foreign language teachers: findings of a European study. *Language Assessment Quarterly*, 11(4), 374–392. <https://doi.org/10.1080/15434303.2014.960046>

- Wach, A. (2012). Classroom-based language efficiency assessment: A challenge for EFL teachers. *Glottodidactica, An International Journal of Applied Linguistics*. 39(1), 81-92. <https://doi.org/10.14746/gl.2012.39.1.8>
- Yan, X., Zhang, C., & Fan, J. J. (2018). Assessment knowledge is important, but ...”: How contextual and experiential factors mediate assessment practice and training needs of language teachers. *System*, 74, 158–168. <https://doi.org/10.1016/j.system.2018.03.003>

## APPENDICES

## Appendix A. Approval from the Members of the Ethics Committee

T.C	
ÇAĞ ÜNİVERSİTESİ	
SOSYAL BİLİMLER ENSTİTÜSÜ	
TEZ / ARAŞTIRMA / ANKET / ÇALIŞMA İZİNİ / ETİK KURULU İZİNİ TALEP FORMU VE ONAY TUTANAK FORMU	
ÖĞRENCİ BİLGİLERİ	
T.C. NOSU	
ADI VE SOYADI	Gamze Kapucu
ÖĞRENCİ NO	2020008011
TEL. NO.	
E - MAİL ADRESLERİ	
ANA BİLİM DALI	İngiliz Dili Eğitimi
HANGİ AŞAMADA OLDUĞU (DERS / TEZ)	Tez
İSTEKTE BULUNDUĞU DÖNEME AİT DÖNEMLİK KAYDININ YAPILIP-YAPILMADIĞI	2022 / 2023 - GÜZ DÖNEMİ KAYDINI YENİLEDİM.
ARAŞTIRMA/ANKET/ÇALIŞMA TALEBİ İLE İLGİLİ BİLGİLER	
TEZİN KONUSU	This research attempts to add to the expanding body of knowledge about conceptualizing LAL by addressing the issue of teacher educators' understanding of language assessment and how applicable Taylor's (2013) LAL model is to Turkish teacher educators.
TEZİN AMACI	This study aims to explore the LAL conceptions of teacher educators in Turkey by making a comparison with Taylor's (2013) LAL model for stakeholders.
TEZİN TÜRKÇE ÖZETİ	Bu çalışma öğretmen eğitimcilerinin Yabancı Dilde Ölçme ve Değerlendirme Okuryazarlığını nasıl kavramsallaştırdıklarını incelemek ve Taylor'ın (2013) LAL modelinin Türk öğretmen eğitimcileri için ne kadar uygulanabilir olduğu konusunu ele alarak LAL'i kavramsallaştırma konusundaki genişleyen bilgi birikimine katkıda bulunmaya çalışmaktadır. Bu çalışma, paydaşlar için Taylor'ın (2013) LAL modeli ile bir karşılaştırma yaparak Türkiye'deki öğretmen eğitimcilerinin LAL anlayışlarını keşfetmeyi amaçlamaktadır.
ARAŞTIRMA YAPILACAK OLAN SEKTÖRLER/ KURUMLARIN ADLARI İZİN ALINACAK OLAN	Çalışma gönüllülük esasına dayanmaktadır. Herhangi bir kuruma bağlı değildir.
İZİN ALINACAK OLAN KURUMA AİT BİLGİLER (KURUMUN ADI- ŞUBESİ/ MÜDÜRLÜĞÜ - İLİ - İLÇESİ) YAPILMAK İSTENEN ÇALIŞMANIN İZİN ALINMAK İSTENEN KURUMUN HANGİ İLÇELERİNE/ HANGİ KURUMUNA/ HANGİ BÖLÜMÜNDE/ HANGİ ALANINA/ HANGİ KONULARDA/ HANGİ GRUBA/ KİMLERE/ NE UYGULANACAĞI GİBİ AYRINTILI BİLGİLER UYGULANACAK OLAN ÇALIŞMAYA AİT ANKETLERİN/ ÖLÇEKLERİN BAŞLIKLARI/ HANGİ ANKETLERİN - ÖLÇELERİN	
EKLER (ANKETLER, ÖLÇEKLER, FORMLAR, ... V.B. GİBİ EVRAKLARIN İSİMLERİYLE BİRLİKTE KAÇ ADET/SAYFA OLDUKLARINA AİT BİLGİLER İLE AYRINTILI YAZILACAKTIR)	1) 3 (üç) Sayfa Yarı-yapılandırılmış Mülakat Soruları 2) 1 (bir) Sayfa Bilgilendirilmiş Onam Formu
ÖĞRENCİNİN ADI - SOYADI: Gamze Kapucu	ÖĞRENCİNİN İMZASI: Enstitü müdürlüğünde evrak aslı ıslak imzalıdır. TARİH: 22 / 11 / 2022
TEZ/ ARAŞTIRMA/ANKET/ÇALIŞMA TALEBİ İLE İLGİLİ DEĞERLENDİRME SONUCU	

TEZ/ ARAŞTIRMA/ANKET/ÇALIŞMA TALEBİ İLE İLGİLİ DEĞERLENDİRME SONUCU						
1. Seçilen konu Bilim ve İş Dünyasına katkı sağlayabilecektir.						
2. Anılan konu ..... faaliyet alanı içerisine girmektedir.						
1.TEZ DANIŞMANININ ONAYI	2.TEZ DANIŞMANININ ONAYI (VARSA)	ANA BİLİM DALI BAŞKANININ ONAYI		SOSYAL BİLİMLER ENSTİTÜSÜ MÜDÜRÜNÜN ONAYI		
Adı - Soyadı: Şehnaz ŞAHİNKARAKAŞ	Adı - Soyadı: .....	Adı - Soyadı: Şehnaz ŞAHİNKARAKAŞ		Adı - Soyadı: Murat KOÇ		
Unvanı: Prof. Dr.	Unvanı: .....	Unvanı: Prof. Dr.		Unvanı: Prof. Dr.		
İmzası: Enstitü müdürlüğünde evrak aslı ıslak imzalıdır.	İmzası: .....	İmzası: Enstitü müdürlüğünde evrak aslı ıslak imzalıdır.		İmzası: Enstitü müdürlüğünde evrak aslı ıslak imzalıdır.		
22.11.2022	..... / ..... / 20....	..... / ..... / 20.....		..... / ..... / 20....		
ETİK KURULU ASIL ÜYELERİNE AİT BİLGİLER						
Adı - Soyadı: Şehnaz ŞAHİNKARAKAŞ	Adı - Soyadı: Yücel ERTEKİN	Adı - Soyadı: Deniz Aynur GÜLER	Adı - Soyadı: Mustafa BAŞARAN	Adı - Soyadı: Mustafa Tevfik ODMAN	Adı - Soyadı: Hüseyin Mahir FİSUNOĞLU	Adı - Soyadı: Jülide İNÖZÜ
Unvanı : Prof. Dr.	Unvanı : Prof. Dr.	Unvanı: Prof. Dr.	Unvanı : Prof. Dr.	Unvanı: Prof. Dr.	Unvanı : Prof. Dr.	Unvanı : Prof. Dr.
İmzası: Enstitü müdürlüğünde evrak aslı ıslak imzalıdır.	İmzası: Enstitü müdürlüğünde evrak aslı ıslak imzalıdır.	İmzası: Enstitü müdürlüğünde evrak aslı ıslak imzalıdır.	İmzası: Enstitü müdürlüğünde evrak aslı ıslak imzalıdır.	İmzası: Enstitü müdürlüğünde evrak aslı ıslak imzalıdır.	İmzası: Enstitü müdürlüğünde evrak aslı ıslak imzalıdır.	İmzası: Enstitü müdürlüğünde evrak aslı ıslak imzalıdır.
..... / ..... / 20.....	..... / ..... / 20....	..... / ..... / 20.....	..... / ..... / 20....	..... / ..... / 20.....	..... / ..... / 20.....	..... / ..... / 20.....
Etik Kurulu Jüri Başkanı - Asıl Üye	Etik Kurulu Jüri Asıl Üyesi	Etik Kurulu Jüri Asıl Üyesi	Etik Kurulu Jüri Asıl Üyesi	Etik Kurulu Jüri Asıl Üyesi	Etik Kurulu Jüri Asıl Üyesi	Etik Kurulu Jüri Asıl Üyesi
OY BİRLİĞİ İLE	<input checked="" type="radio"/>	Çalışma yapılacak olan tez için uygulayacak olduğu Anketleri/Formları/Ölçekleri Çağ Üniversitesi Etik Kurulu Asıl Jüri Üyelerince İncelenmiş olup, ..... / ..... / 20..... - ..... / ..... / 20..... tarihleri arasında uygulanmak üzere gerekli iznin verilmesi tarafımızca uygundur.				
OY ÇOKLUĞU İLE	<input type="radio"/>					
AÇIKLAMA: BU FORM ÖĞRENCİLER TARAFINDAN HAZIRLANDIKTAN SONRA ENSTİTÜ MÜDÜRLÜĞÜ SEKRETERLİĞİNE ONAYLAR ALINMAK ÜZERE TESLİM EDİLECEKTİR. AYRICA FORMDAKİ YAZI ON İKİ PUNTO OLACAK ŞEKİLDE YAZILACAKTIR.						



## Appendix B. Consent Form

**ÇAĞ ÜNİVERSİTESİ  
SOSYAL BİLİMLER ENSTİTÜSÜ  
ETİK KURULU**

### BİLGİLENDİRİLMİŞ ONAM FORMU

Bu formun amacı araştırma ile ilgili olarak sizi bilgilendirmek ve katılmanız ile ilgili izin almaktır.

Bu kapsamda "Teacher Educators' Conceptions on Language Assessment Literacy: A Case Study from Türkiye" başlıklı araştırma Gamze KAPUCU tarafından **gönüllü katılımcılarla** yürütülmektedir. Araştırma sırasında sizden alınacak bilgiler gizli tutulacak ve sadece araştırma amaçlı kullanılacaktır. Araştırma sürecinde konu ile ilgili her türlü soru ve görüşleriniz için aşağıda iletişim bilgisi bulunan araştırmacıyla görüşebilirsiniz. Bu araştırmaya **katılmama** hakkınız bulunmaktadır. Aynı zamanda çalışmaya katıldıktan sonra çalışmadan **çıkabilirsiniz**. Bu formu onaylamanız, **araştırmaya katılım için onam verdiğiniz** anlamına gelecektir.

#### Araştırmayla İlgili Bilgiler:

**Araştırmanın Amacı:** Bu çalışma, Türkiye'deki öğretmen eğitimcilerinin Yabancı Dilde Ölçme ve Değerlendirme Okuryazarlığını nasıl kavramsallaştırdıklarını araştırmayı amaçlamaktadır.

**Araştırmanın Nedeni:** Bu araştırma yüksek lisans tez çalışması kapsamında yürütülmektedir.

**Araştırmanın Yürütüleceği Yer:** Eskişehir, Isparta, Hatay, Afyon, Nevşehir

#### Çalışmaya Katılım Onayı:

Araştırmanın amacını, nedenini, yürütüleceği yer ile ilgili bilgileri okudum ve gönüllü olarak üzerime düşen sorumlulukları anladım. Araştırma ile ilgili ayrıntılı açıklamalar yazılı ve sözlü olarak tarafıma sunuldu. Bu araştırma ile ilgili faydalar ve riskler ile ilgili bilgilendirildim.

Bu araştırmaya kendi isteğimle, hiçbir baskı ve zorlama olmaksızın katılmayı kabul ediyorum.

#### Katılımcının (Islak imzası ile<sup>\*\*\*</sup>)

Adı-Soyadı:  
İmzası<sup>\*\*\*</sup>:

#### Araştırmacının

Adı-Soyadı:  
e-posta:

İmzası: Evrak Aslı Enstitü Müdürlüğünde ıslak imzalıdır.

## Appendix C. Semi-Structured Interview Guide for Teacher Educators

### Semi-Structured Interview Guide (Teacher Educators)

1. Can you please state your age, if you don't mind?
2. What is your area of expertise?
3. How long have you been working as a teacher educator?
4. Have you ever taught any assessment courses or taught the subject of assessment as a teacher educator? What did these courses cover?
5. Have you been involved in language assessments in any other capacity (e.g., grading/advising/research)?
6. Have you published any articles, books, reports (or other types of publications) on assessment and evaluation (peer-reviewed or non-peer-reviewed)? If possible, could you mention what they are?
7. Do you have any experience in teaching and assessing English as a foreign language in English language classrooms (in Turkey or elsewhere)?
8. How knowledgeable do you consider yourself in assessment and evaluation in English? On a scale of 1 to 6, how would you rate yourself? / What score would you give yourself out of 6?

### Language Assessment Literacy

#### General Questions

9. What kind of knowledge and skills does an English teacher working in the Ministry of Education (MEB) need in order to conduct effective assessment and evaluation?
10. Regarding your answer to question 9, do you think there are any skills or areas that are more important than others? Do you consider them all equally important?
11. How important do you think assessment and evaluation are in teaching and learning English? Please rate on a scale of 1 to 6 and explain why.

#### Specific Questions

To what extent are the following aspects important for the assessment competence of an English teacher? Please rate the following statements on a scale of 1 to 6, where 1 represents the lowest importance, and 6 represents the highest importance:

12. It is important to have knowledge about areas where assessment can support learning (formative assessment, curriculum teaching & learning processes).
13. It is important to have knowledge about self-assessment.
14. It is important to have knowledge about peer assessment.
15. It is important to have knowledge about the following:
  - a. Language testing methods and how they can be appropriately used in English classes, e.g., multiple-choice tests.
  - b. How to construct good questions.

- c. Statistical measurements (e.g., mean, standard deviation, measurement error) for interpreting the results of multiple-choice tests.
16. Please rate the following on a scale of 1 to 6:
    - a. The importance of having knowledge about process assessment.
    - b. The importance of having knowledge about outcome/product assessment.
  17. It is important to have knowledge about different types of assessment.
  18. It is important to have knowledge about how assessment can be used as a diagnostic tool.
  19. It is important to have knowledge about different assessment purposes (diagnostic, ranking, promoting learning).
  20. It is important to have knowledge about large-scale tests such as national tests (YDS, YÖKDİL), international tests (TOEFL, IELTS), PISA, etc.
  21. It is important to have knowledge about how assessment can be used to motivate students.
  22. It is important to have knowledge about:
    - a. Second language acquisition theories.
    - b. Communicative competence theories.
  23. It is important to have knowledge about common language criteria such as CEFR (Common European Framework of Reference for Languages).
  24. It is important to have knowledge about the national education program (general curriculum and English language teaching programs).
  25. It is important to have knowledge about the National Education Law and other regulations.
  26. It is important to have knowledge about contextual issues related to assessment (e.g., assessment criteria of education authorities, how teachers in your region evaluate student performance, etc.).
  27. It is important to have knowledge about assessment/testing theories (e.g., validity, reliability, practicality, etc.).
  28. It is important to have knowledge about ethical issues (e.g., transparency, unintended use of assessment results, etc.).
  29. It is important to have knowledge about how grading scales/scoring rubrics should be used.
  30. It is important to have knowledge about how assessment results should be appropriately communicated (e.g., how to explain national test results to parents).
  31. It is important to have knowledge about how societal values can influence assessment.
  32. It is important to have knowledge about how one's own views on assessment (values, perspectives) can influence assessment.
  33. It is important to have knowledge about the history of language assessment.

## Appendix D. Closed-ended Questionnaire Form

8. On a scale from one to six, how knowledgeable do you consider yourself to be when it comes to the topic 'language assessment'?

11. On a scale from one to six, how important would you say that assessment is in English teaching and learning?

*To what extent are the following aspects important for an English teacher's assessment competence? On a scale from one to six, where one is the lowest, and six is the highest, please judge the following questions/statements:*

12. It is important to have knowledge of how assessment can promote learning 1 2 3 4 5 6

13. It is important to have knowledge of self-assessment 1 2 3 4 5 6

14. It is important to have knowledge of peer-assessment 1 2 3 4 5 6

15. It is important to have knowledge of:

a. how language testing can be used appropriately in the English

language classroom, e.g., multiple choice 1 2 3 4 5 6

b. how good items are created. 1 2 3 4 5 6

c. statistical measures in order to interpret results from multiple-choice tests (e.g., mean, standard deviation, measurement error) 1 2 3 4 5 6

16. How do you understand the terms 'formative' and 'summative' assessment? On a scale from one to six, score the following statements:

a. It is important to have knowledge of formative assessment 1 2 3 4 5 6

b. It is important to have knowledge of summative assessment 1 2 3 4 5 6

17. It is important to have knowledge of different types of assessment 1 2 3 4 5 6

18. It is important to have knowledge of how assessment can be used as a diagnostic tool. 1 2 3 4 5 6

19. It is important to have knowledge of different purposes of assessment (diagnostic, ranking, promoting learning) 1 2 3 4 5 6

20. It is important to have knowledge of large-scale testing, such as national tests, and the English Proficiency Exam for Academic Purposes (YDS) and Higher Education Institutions Foreign Language Test (YÖKDİL), etc. 1 2 3 4 5 6

21. It is important to have knowledge of how assessment can be used to motivate students. 1 2 3 4 5 6

22. a. It is important to have knowledge of

a. languages and language learning 1 2 3 4 5 6

b. theories of communicative competence 1 2 3 4 5 6

23. It is important to have knowledge of language frameworks, such as CEFR. If yes, why? What benefits does knowledge of the CEFR provide to the teacher? 1 2 3 4 5 6

24. It is important to have knowledge of the national curriculum (the general part and the English subject curriculum) 1 2 3 4 5 6
25. It is important to have knowledge of Education Act and The Basic Law of National Education 1 2 3 4 5 6
26. It is important to have knowledge of local issues which may be relevant for assessment (e.g. the local educational authorities' assessment criteria, how teachers in your local district assess student performance etc.) 1 2 3 4 5 6
27. It is important to have knowledge of assessment/testing theory (related to «validity», «reliability» etc.) 1 2 3 4 5 6
28. It is important to have knowledge of ethical issues (fairness, use of assessment results for purposes other than what was intended etc.) 1 2 3 4 5 6
29. It is important to have knowledge of how to use rating scales / scoring rubrics 1 2 3 4 5 6
30. It is important to have knowledge of how to communicate assessment results in appropriate ways (e.g. how to explain to parents the results from national tests) 1 2 3 4 5 6
31. It is important to have knowledge of how values in society may affect assessment 1 2 3 4 5 6
32. It is important to have knowledge of how one's own view on assessment (values, outlook) may affect assessment 1 2 3 4 5 6
33. It is important to have knowledge of the history of language assessment 1 2 3 4 5 6

## Appendix E. Scientific Research and Publication Ethics Board Decision



T.C.  
ÇAĞ ÜNİVERSİTESİ  
Sosyal Bilimler Enstitüsü

Sayı : E-23867972-050.01.04-2200008855

22.11.2022

Konu : Bilimsel Araştırma ve Yayın Etiği  
Kurulu Kararı Alınması Hk.

REKTÖRLÜK MAKAMINA

**İlgi:** 09.03.2021 tarih ve E-81570533-050.01.01-2100001828 sayılı Bilimsel Araştırma ve Yayın Etiği Kurulu konulu yazınız.

İlgi tarihli yazınız kapsamında Üniversitemiz Sosyal Bilimler Enstitüsü bünyesindeki Lisansüstü Programlarda halen tez aşamasında kayıtlı olan **Gamze Kapucu** isimli öğrencimize ait tez evraklarının "Üniversitemiz Bilimsel Araştırma ve Yayın Etiği Kurulu Onayları" alınmak üzere Ek'te sunulmuş olduğunu arz ederim.

Prof. Dr. Murat KOÇ  
Sosyal Bilimler Enstitüsü Müdürü

Ek : 1 adet öğrenciye ait tez evrakları dosyası.

## Appendix F. Ethical Permission Letter



T.C.  
ÇAĞ ÜNİVERSİTESİ  
Rektörlük

Sayı : E-81570533-044-2200009196  
Konu : Bilimsel Araştırma ve Yayın Etiği  
Kurul İzni Hk.

06.12.2022

### SOSYAL BİLİMLER ENSTİTÜSÜ MÜDÜRLÜĞÜNE

İlgi : a) 22.11.2022 tarih ve E-23867972- 050.01.04-2200008854 sayılı yazınız.  
b) 22.11.2022 tarih ve E-23867972- 050.01.04-2200008855 sayılı yazınız.  
c) 23.11.2022 tarih ve E-23867972- 050.01.04-2200008859 sayılı yazınız.  
ç) 01.12.2022 tarih ve E-23867972- 050.01.04-2200009060 sayılı yazınız.

İlgi yazılarda söz konusu edilen **Muhammed Alperen KILIÇ, Mehmet Sıtkı TEMUR, Özen TEKİN, Mustafa Recep OKTAY, Hülya YARDIMCI, Nazlı BAYRAM, Gamze KAPUCU ve Dila BARAN** isimli öğrencilerimize ait tez evrakları Bilimsel Araştırma ve Yayın Etiği Kurulunda incelenerek uygun görülmüştür.

Bilgilerinizi ve gereğini rica ederim.

Prof. Dr. Ünal AY  
Rektör