BASKENT UNIVERSITY

INSTITUTE OF EDUCATIONAL SCIENCES

DEPARTMENT OF FOREIGN LANGUAGES EDUCATION

MASTER'S DEGREE PROGRAMME IN ENGLISH LANGUAGE TEACHING WITH THESIS

EXPLORING THE RELATIONSHIP BETWEEN EFL TEACHERS' LANGUAGE ASSESSMENT LITERACY AND THEIR REFLECTIVE TEACHING

PREPARED BY

MERVE ARICI

MASTER THESIS

ANKARA – 2022

BASKENT UNIVERSITY

INSTITUTE OF EDUCATIONAL SCIENCES

DEPARTMENT OF FOREIGN LANGUAGES EDUCATION

MASTER'S DEGREE PROGRAMME IN ENGLISH LANGUAGE TEACHING WITH THESIS

EXPLORING THE RELATIONSHIP BETWEEN EFL TEACHERS' LANGUAGE ASSESSMENT LITERACY AND THEIR REFLECTIVE TEACHING

PREPARED BY
MERVE ARICI

MASTER THESIS

THESIS SUPERVISOR
ASSOC.PROF.DR. FARHAD GHORBANDORDINEJAD

ANKARA – 2022

BAŞKENT ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Yabancı Diller Eğitimi Anabilim Dalı İngiliz Dili Eğitimi Tezli Yüksek Lisans Programı çerçevesinde Merve ARICI tarafından hazırlanan bu çalışma, aşağıdaki jüri tarafından Yüksek Lisans Tezi olarak kabul edilmiştir.

Tez Savunma Tarihi: 27/07/ 2022

Tez Adı: Exploring the Relationship between EFL Teachers' Language Assessment Literacy and Their Reflective Teaching

Tez Jüri Üyeleri (Unvanı, Adı - Soyadı, Kurumu)	**İmza**
Doç. Dr. Farhad Ghorbandordinejad	……………
Prof. Dr. Abdulvahit Çakır	……………
Doç. Dr. Cemal Çakır	……………
……………………………………………	……………
……………………………………………	…………… Gerekli Durumda
……………………………………………	…………… Gerekli Durumda

ONAY

………………………………………………….

…………………………….... Enstitüsü Müdürü

Tarih: …/ …/ ……..

BAŞKENT ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ
YÜKSEK LİSANS TEZ ÇALIŞMASI ORİJİNALLİK RAPORU

Tarih: 27/ 06/ 2022

Öğrencinin Adı, Soyadı: Merve ARICI

Öğrencinin Numarası:21810400

Anabilim Dalı: İngiliz Dili Öğretimi.

Programı: İngiliz Dili Öğretimi Tezli Yüksek Lisans Programı.

Danışmanın Unvanı/Adı, Soyadı: Dr. Öğr. Üyesi Farhad Ghorbandordinejad

Tez Başlığı: İngilizce Öğretmenlerinin Dil Değerlendirme Okuryazarlıkları ile Yansıtıcı Öğretimleri Arasındaki İlişkinin Keşfedilmesi

Yukarıda başlığı belirtilen Yüksek Lisans/Doktora tez çalışmamın; Giriş, Ana Bölümler ve Sonuç Bölümünden oluşan, toplam 68. sayfalık kısmına ilişkin, 27 / 06 / 2022 tarihinde şahsım/tez danışmanım tarafından Ithenticate adlı intihal tespit programından aşağıda belirtilen filtrelemeler uygulanarak alınmış olan orijinallik raporuna göre, tezimin benzerlik oranı % 19 .'dır. Uygulanan filtrelemeler:

1. Kaynakça hariç

2. Alıntılar hariç

3. Beş (5) kelimeden daha az örtüşme içeren metin kısımları hariç

"Başkent Üniversitesi Enstitüleri Tez Çalışması Orijinallik Raporu Alınması ve Kullanılması Usul ve Esaslarını" inceledim ve bu uygulama esaslarında belirtilen azami benzerlik oranlarına tez çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

Öğrenci İmzası:…………………..

ONAY

Tarih: 27 / 06 / 2022 .

Öğrenci Danışmanı Unvan, Ad, Soyad, İmza:

Dr. Öğr. Üyesi Farhad Ghorbandordinejad

To my family and my dear brother who always believed in me and to my dearest husband who supported me about everything without whom, this thesis would have not been possible, and to my twins waiting for me in heaven

ACKNOWLEDGEMENTS

First, I would like to express my deepest appreciation to my supervisor Assoc. Prof. Dr. Farhad GHORBANDORDINEJAD, for being a marvelous mentor to me and for all his support throughout my thesis process. His bits of help and pieces of advice are invaluable for my master's degree and thesis.

I would like to thank Assoc. Prof. Dr. Senem ÜSTÜN KAYA, and my lecturers for their effort throughout my degree, my master's degree, and my thesis process.

Last but not the least, I would like to thank my loving, and supportive husband Zahid and my heroic, angel mother, my dearest, clever brother, and my generous father. Words cannot express how thankful I am to have you in my life. Your constant support encouraged me throughout my thesis study. In addition, I would like to thank Büşra CİVİL and my other valuable friends for their friendship that allowed me to conduct this thesis.

ÖZET

Merve ARICI

İngilizce Öğretmenlerinin Dil Değerlendirme Okuryazarlıkları ile Yansıtıcı Öğretimleri Arasındaki İlişkinin Keşfedilmesi, Başkent Üniversitesi Eğitim Bilimleri Enstitüsü, Yabancı Diller Anabilim Dalı İngiliz Dili Öğretimi Tezli Yüksek Lisans Programı

2022

Pek çok dil öğretmeni, dil değerlendirmesindeki araştırma-uygulama boşluğunu kapatmak için mücadele ederken, öğretmen eğitimi bulmacasının önemli bir parçasını henüz yerine oturtmuş değiller. Sınıf eğitiminin yanı sıra, öğretmenlerin mesleki gelişimi (MG), yansıtma için fırsatlar içermelidir. Bu çalışma, EFL öğretmenlerinin dil değerlendirme okuryazarlığı ile Türk İngilizce öğretmenlerinin yansıtıcı öğretimi arasındaki ilişkiyi araştırıyor, mevcut sınıf değerlendirme uygulamalarını yeniden inceliyor ve yakın tarihli değerlendirme araştırmalarının pratik geçerliliğini yansıtarak gelecekteki dil değerlendirme eylemlerine yönelik planlarını araştırıyor. Araştırmaya dayalı teoriler ve dil değerlendirme prosedürleri üzerine öğretmen yansıması uyumludur ve bulgulara göre öğretmenlerin gelecekteki dil değerlendirme performansına yönelik yansımalarına rehberlik eder (eylem üzerine yansıma). Ayrıca öğretmenleri, dil değerlendirme okuryazarlıklarını pratik olarak geliştirmek için önerilerde bulunmaya teşvik ettik. Veri toplama yöntemi, 100 İngilizce öğretmeni arasında dağıtılan online anketi içerir ve toplanan veriler, çok aşamalı bir sistematik içerik analizinin ardından SPSS yazılımı kullanılarak analiz edildi.

Anahtar Sözcükler: Dil değerlendirme okuryazarlığı, Yansıtıcı öğretim, Türk İngilizce öğretmenleri

ABSTRACT

Merve ARICI

Exploring the Relationship between EFL Teachers' Language Assessment Literacy and Their Reflective Teaching, Başkent University Institute of Educational Sciences, Department of Foreign Languages Master in English Language Teaching

2022

As many language teachers struggle to bridge the research-practice gap in language evaluation, they have not yet put a crucial piece of the puzzle of teacher education in place. Besides classroom instruction, teachers' Professional Development (PD) should include opportunities for reflection. This study explores the relationship between EFL teachers' language assessment literacy and their reflective teaching of Turkish EFL teachers, re-examining their current classroom assessment practices and their plans for future language assessment actions by reflecting on the practical relevance of recent assessment research. Teacher reflection on research-based theories and language assessment procedures are compatible and guides teachers' reflection on future language assessment performance, according to the findings (reflection-on-action). We also encouraged teachers to make suggestions for practically improving their Language Assessment Literacy (LAL). The data collection method includes a questionnaire distributed among 100 EFL teachers, and the collected data was analyzed using SPSS software following a multiple-phased systematic content analysis.

Keywords: Language assessment literacy, Reflective teaching,Turkish EFL instructors

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATION LIST

AaL	Assessment as Learning
AfL	Assessment for Learning
AoL	Assessment of Learning
EFL	English as a Foreign Language
ELT	English Language Teaching
ELTE	English Language Evaluation and Testing
ESL	English as a Second Language
LAL	Language Assessment Literacy
LTA	Language Teaching Assessment
MC	Multiple Choice
TAL	Teachers' Assessment Literacy
WAL	Writing Assessment Literacy

1. INTRODUCTION

In two distinct institutions in Turkey (Gazi University and Başkent University), this study investigates the relationship between EFL teachers' language assessment literacy and their reflective teaching. The background of the study, statement of the problem, purpose of the study, research questions, the significance of the study, limitations of the study, and lastly definitions of key terms are all included in the study's introduction section.

1.1. The Background of the Study

Assessing student competence and performance, "interpreting results from these assessments, and using these findings to improve student learning or program effectiveness" has been described as the instructors' assessment literacy (Webb, 2002, p. 1). There are a variety of pedagogical strategies that can be used in the classroom to encourage learning through self and peer assessment, as well as continual descriptive feedback and the establishment of assessment standards.

Teachers that are knowledgeable in science are better equipped to select the most effective assessment techniques, provide valid assessment activities, provide feedback to students, and evaluate the teaching and learning process. They also have a thorough understanding of the theoretical and philosophical aspects of assessing students' learning.

An assessment literacy that covers not only what and how to assess but also instructors' knowledge, beliefs, and behaviors have recently been advocated by Crusan et al. (2016). Several variables at play in a teacher's decision-making process can either encourage or hinder student learning (Crusan et al., 2016; Med & Atay, 2017; Weigle, 2007; Weigle et al., 2016 White, 2009).

Students' time, motivation, and self-esteem can be negatively impacted by teachers' lack of evaluation knowledge and experience as well as ineffective assessment methods (Crusan et al., 2016). Teachers spend a large portion of their professional time on assessment-related tasks, hence they frequently struggle to execute it well (Crusan et al.,

2016; Fulcher, 2012; Jin, 2010; Lam, 2015; Lan & Fan, 2019; Mertler, 2009; Stiggins, 1999, 2014; Tsagari & Vogt, 2017; White, 2009; Zhu, 2004).

As for the writing talent, teacher educators of that in their classes because they have not (Pytlikzillig & Bruning, 2009). For this reason, pre-and in-service EFL teachers must receive assessment training to evaluate their students' progress and maximize the assessment's learning potential (Crusan et al., 2016, p. 46). Besides the instructional situation and the initial language acquisition of the teachers (Boyles, 2006; Hirvela & Belcher, 2007; Malone, 2013; Taylor, 2009; Vogt & Tsagari, 2014; Weigle, 2007; White, 2009).

A few research has examined the writing assessment knowledge, beliefs, and training needs of second language (L2) teachers. ESL and EFL writing teachers from 41 nations on five continents were investigated by Crusan et al. (2016), one pioneer in this field. Teacher backgrounds and experience varied widely, but on average, 26% of teachers had received little or no training in teaching and assessing written communication.

While Crusan et al. (2016) looked into these concerns; the current study only looks at the Turkish setting because the subject has not been examined here yet. Soltanpour and Valizadeh (2019) assessed Iranian EFL instructors' writing assessment literacy to determine their training needs. As also investigated what the teachers believe about the accuracy of scoring in writing assessments, as well as assessment issues in general in writing classrooms and various assessment methods for writing.

Because of the use of post-method technology, reflective teaching has become more critical in EFL teaching. For educational systems, the most important components are post-method instructors who are self-reliant, self-confident individuals (Marcos, et al., 2011). They are highly esteemed individuals whose contributions to the classroom are priceless. They are aware of their own issues and less reliant on outside (external) expert knowledge for assistance (Cousin, 2000). It is said that they "theorize what they do and do what they theorize" (Kumaravadivelu, 1994, p. 30).

One of the most important concepts to emerge in post-technological pedagogy has been the concept of reflection, which is the subject of numerous studies from different perspectives and dimensions. It has been argued that (Akbari 2007; Farrah Farrell 2012; Farrell 2016; Marcos Sanchez & Tillema, 2011; Akbari, 2007). A method called "reflective

teaching" is based on the idea that teachers can improve their own perceptions of teaching and their level of satisfaction with their own teaching by critically reflecting on their overall performance. The reflective approach to teaching in teacher education programs aims to improve the ability of teachers to analyze the teaching process in order to prepare them for powerful mastery.

Schon proposed one of the most widely used conceptualizations of reflective teaching (1996). Schon believes that reflective teaching can be useful for new teachers who want to see how much they can learn from more experienced colleagues (Akbari, 2005). Using expert instructors to supervise new instructors ensures they do not forget they are practicing using expertise cautiously (Ferraro, 2000). Reflection-for-action, reflection-in-action, and reflection-on-action are three of Schon's (1996) recommended aspects of reflective teaching. Instructors use reflection-for-action to prepare for issues or scenarios that they expect to arise in their classes in the future Burhan-Horasanli and Ortaçtepe (2016). When you are doing something, it is important to stop and think about what you are doing. Once the mastering session has been meticulously planned and prepared, you will want to monitor the session's progress at all times. Because you are focused, you are better able to adapt to changing circumstances and "think on your feet" (Scales, 2008, p. 11). Reflection-on-action is a reflection after you have completed the task. After the consultation is over, you will be able to replicate, examine, and evaluate the learning and teaching. Post-method reflection informs your next planning and practice, resulting in a cycle of "perseverance with" (Scales, 2008, p. 11).

1.2. Statement of the Problem

The issues surrounding Turkish EFL teachers' writing assessment literacy (henceforth, WAL), beliefs, and training requirements merit consideration because figuring out what Turkish EFL teachers already know and believe, as well as what writing assessment training requirements they have, could pave the way for offering support for more in-depth study on this subject. To the best of the researcher's knowledge, there have been few studies that have comprehensively investigated the aforementioned difficulties in Turkey. using pre-made tests in the classroom, giving comments on assessments, and using peer or self evaluation. The current study, like Mede and Atay (2017), examined the Turkish EFL teachers' prior assessment training experiences, their perceived needs for training in this area, as well as

their attitudes toward testing and assessment practices in language programs. However, the present study's focus was on testing and assessment in L2 writing rather than the general field of language testing and assessment. To put it briefly, it was felt that there was a particular need to look at the writing assessment literacy, beliefs, and training needs of Turkish EFL teachers in order to inform the authorities, academics, and teacher educators in Turkey.This study finds out how much training EFL instructors have had in language assessment and reflective teaching, theoretically and practically, in teacher training and whether they see a need for further training in this area. In contrast to other studies in the field, this one will not only look at the current state of pre-service language teachers' assessment literacy but also at the large extent to which evaluation knowledge is implemented in a practicum experience and that assessment theory and practice are covered in education experience courses and practicum. The study will also look into how they are being taught in school experience and practicum courses. In addition, the research will look into the impact of the English Language Evaluation and Testing (ELTE) course on students' schoolwork and internships.

1.3. Purpose of the Study

The main purpose of this study is to explore the relationship between English as a Foreign Language (EFL) teachers' language assessment literacy and their reflective teaching. The researcher wants to see if reflective teaching in their classrooms predicted teachers' assessment literacy. As a result, the researcher puts the following research questions out for discussion:

1.4. Research questions

The research highlights the following main research questions: Is there a relationship between EFL teachers' language assessment literacy and their reflective teaching?
1:How proficient do EFL teachers consider themselves writing instructors?

2:What do EFL teachers think of scoring accuracy when it comes to writing assessment?

3:What do EFL teachers think of general assessment when it comes to writing proficiency?

4:What do EFL teachers think of different writing assessment methods?

1.5. The Significance of the Study

As a complex and demanding profession, teaching requires qualified teachers. The success of every educational system depends heavily on the quality of its teachers. In language learning and teaching, the instructor plays an important role, as noted by Campbell (2000). When it comes to teaching, many studies have shown that students assessed in the classroom performed better academically than those not (Earl & Katz, 2006). Consequently, teachers are expected to evaluate their pupils' progress as a fundamental component of their work (Brumfit & Johnson, 1979). Even though English teachers should be familiar with alternative assessment techniques, many of them still use the old-fashioned approach. As a result, an increasing number of teacher and administrator professional development programs have focused on assessment literacy, and numerous scholars and organizations have identified specific curriculum areas in which instructors must enhance their assessment skills.

John Dewey first proposed retrospective thinking in 1933, and it has since been extensively investigated. Dewey outlined the concept of reflective thought, defining it as the careful consideration of any idea in light of previous knowledge and future aims. Teachers and student teachers can use the information they gather through their observations, attitudes, ideas, and teaching practices as a basis for critical reflection on teaching, according to Richards & Lockhart (1996). (p.1). Reflective practice is a way of approaching teaching and teacher education founded on the idea that teachers may better their teaching by critically reflecting on their own classroom experiences. A reflective teacher examines the genesis, principles, and outcomes of their students' work in the classroom. Teaching becomes more accessible for teachers when they engage in these in-depth talks that focus on specific practices (Rose, 2007).

To paraphrase Griffiths, "professionals are confronted with conditions which they perceive as unique or contain a constituent of surprise," and this is when they engage in reflective action. Professionals do not just apply a theory or previous experience directly (p. 542). Schon1987 (cited in Leitch and Day, 2000) describes a teacher's moral concerns and retrospective analysis of their actions to gain knowledge from experience as "reflection-on-action".

The correlation between reflective teaching and teachers' evaluation literacy has been the subject of theoretical studies, but empirical research has been few. The current study attempts to delve into the association between reflective teaching and EFL teachers' assessment literacy to fill the void left by previous studies in this area.

1.6. Limitations of the Study

1. The sampling of the study will be limited to two universities in Ankara in Turkey. In this respect, the results might not be generalized to the language assessment literacy degree of the EFL teachers.

2. Data will be collected through questionnaires to be administered to the EFL teachers. Therefore, the study's conclusions will be confined to those acquired from these two assessment techniques. Second, because this is a quantitative study, there are a lot more numbers than descriptions. However, to have a deeper and more detailed picture of language assessment practices and skills of the teachers, one needs to observe and analyze the formative and summative assessment instruments used in their classrooms.

Due to its unique nature, there will be certain limitations in conducting any research. During this research, the researcher also encountered some limitations. These limitations can affect the work process and, even to some extent, the study results. The following are some of the limitations and obstacles the researcher has encountered during the research process. Since social sciences, humanities, and management are related to the study and understanding of human activities, and due to the complexity of human behaviour and activities, it is always difficult to collect information from individuals because of their impact. Factors beyond the researcher's control have been suggested. This problem is compounded in our society, where the spirit of research is not institutionalised.

Because the present study focuses on a specific community, it is impossible to generalise data specific to a limited and specific community to other communities. Our research is on "Exploring The Relationship Between EFL Teachers 'Language Assessment Literacy and Their Reflective Teaching". The results of this method are positive in our research, but since we have done this research only on one community, we cannot estimate the results for other communities in the same way.

Since other variables also affect the research. Moreover, in this study, only a limited number of these variables have been addressed.

1.6.1. Limitations under the control of the researcher

- Lack of access to more people.
- This research has been done in a certain period, and this case is one of the significant limitations.
- Due to practical constraints such as time, budget, and lack of research tools or access to the target population.
- Lack of cost and time to thoroughly review all the criteria affecting the research topic.
- Restrictions may include issues such as withdrawal.
- Limitation or uncertainty and validity of data for some measurement measures performed.

1.6.2. Limitations beyond the control of the researcher

- The data collection tool in this study was a questionnaire, which has an inherent limitation.
- There was no similar model for research
- The limitations of research related to the possibility of mental responses of individuals.
- Probability that some experts are unfamiliar with the techniques used in research.
- The novelty of the research subject and, as a result, the lack of internal or external research in the field of the studied subject.

1.7. Definitions of Key Terms

Assessing students' learning abilities, supporting teachers in monitoring student progress, grading students' learning, and assigning students to groups based on institutional standards are all roles of assessment, according to Heaton (1990). According to Stiggins (1992), the purpose of assessment is for teachers to reflect on their teaching, collect relevant evidence, and use that evidence to enhance their teaching.

Language assessment literacy has its roots in educational assessment literacy (Deluca & Klinger, 2010; Stiggins, 1991; Willis, Adie, & Klenowski, 2013). Because a lack of assessment knowledge can "cripple" education, it is clear that assessment knowledge is essential (Popham, 2009, p. 4).

Many academics, including McMillan (2000), Popham (2004), and Stiggins (1999), place assessment literacy at the heart of teaching/learning quality by relating the principles of assessment to what and when to employ assessment methods to collect reliable information/data on students' achievements (Chan, 2008).

Nevertheless, in the analysis of language teaching, studies have proved assessment literacy in the teacher's and testing perspective; familiarity and ability of educators required to plan assessments, administer them, understand and apply the results and a recommendation for teachers in scoring, grading and making judgments on students' academic performance (Boyles 2005; Stiggins 1999; Stoynoff & Chapelle, 2005). (p.1). One of the significant changes in LAL's definition is that it shifts its focus from a teacher and a student to the same thing: learning.

Language teachers are capable of asking and answering tough questions about the purpose of the assessment, the fit of the tool that can be used, the testing conditions, and what will happen based on the results, as demonstrated by Inbar-Lourie (2008), a self-reflective process for language teachers.

Although the concept of language assessment literacy has traditionally been limited to teachers' knowledge of language testing, it has recently been extended to include multiple stakeholder groups (Inbar-Lourie, 2008, McNamara & Roever, 2006). (e.g. policymakers, examination boards, parents, and the general public). For the public to comprehend how assessment instruments and their outputs are used in society, Taylor (2009) underlines the importance of assessment literacy beyond those in test development and research, language teachers, and language teachers.

To better grasp LAL, teachers must be able to integrate their knowledge into historical and social contexts, as well as political and philosophical ones (Fulcher, 2012).

It was found that the ability to evaluate, analyze and utilize student performance data to improve education was considered a sort of language assessment literacy (Falsgraf, 2005).

Much of the study on LAL, on the other hand, described it as a static concept. An optimistic sign for LAL is developing the idea in a larger context, such as its role in self-reflection, the importance of LAL to diverse stakeholder groups, and its incorporation into historical and social contexts. In addition, the LAL conceptual framework needs to include an active and practical perspective on LAL concepts, such as the professional competency aspect highlighted by Taylor (2013).

2. LITERATURE REVIEW

The two key aspects of this investigation are covered in this section of the study. These two key aspects that will be studied and scrutinized are reflective teaching and language assessment literacy, and they have also subtitles. Language assessment literacy includes teachers' assessment literacy, knowledge, conceptions, and classroom-based practices in teacher assessment literacy, teachers' assessment literacy conceptual framework, teachers' assessment literacy on boosting learners' autonomy, assessment literacy skills for EFL teachers, quality of EFL testing, the frequently overlooked secondary EFL tests, literacy testing for language proficiency, language assessment generalities in education. Reflective teaching includes forms and levels of reflection, the importance of teachers' views, the advantages of reflection, reflective teaching adaptation, reflective teaching cycle, reflective teaching assessment, and reflective teaching practices on students' academic achievement, and their comparison.

2.1. Language Assessment Literacy

Language assessment literacy is typically understood as a set of skills, understanding of assessment techniques, and application of proper instruments at the right time that enable a person to comprehend, evaluate, create language tests, and analyze test results (Inbar-Lourie, 2008; Pill & Harding, 2013; Stiggins, 1999). A "skills + knowledge" approach to evaluating literacy was proposed by Davies (2008). While "knowledge" refers to the "relevant background in measuring and linguistic description," "skills" refer to the actual know-how in assessment and creation (p. 328). As is clear from the literature, there has been a shift in language evaluation literacy from a more componential to a developmental perspective (e.g., Brindley, 2001; Davies, 2008; Inbar-Lourie, 2008). Fulcher (2012), for instance, proposed classifying language assessment literacy into three categories: (a) practical knowledge; (b) theoretical and procedural knowledge; and (c) socio-historical understanding. The foundation and primary factor in language evaluation literacy, according to Fulcher, is practical knowledge. Pill and Harding (2013) divided language assessment literacy into four categories: "multidimensional language assessment literacy," "nominal

literacy," "functional literacy," and "procedural and conceptual literacy," with a focus on mathematics and science (p.383).

2.1.1. Teachers' assessment literacy

Using the term "assessment literacy" is a way of describing the knowledge and abilities teachers need to effectively plan for, administer, comprehend, and apply the results of assessments. The concept of perfect assessment has been depicted as an example of assessment literacy (Popham, 2004; Stiggins, 2002). Teachers who are well-versed in this area are better positioned to apply suitable teaching methods and integrate assessment into their lessons (McMillan, 2000). After that, researchers found that instructors' performance in this area was generally inconsistent with what was considered best practice (Galluzzo, 2005; Mertler, 2004). Pre-service teacher candidates are similarly affected by this result since they are more likely to engage in questionable assessment and evaluation methods (Bachor & Baer, 2001; Campbell & Evans, 2000). Teachers' educational programs and career development perspectives are essential in providing teachers with a present understanding of teaching and assessment, notably the expertise to develop formative assessments that elicit students' analytical abilities or assess their improvement and development toward competence (Cizek, 2000).

Researchers have found that many instructors are ill-prepared to administer, interpret, or design diverse assessment methods. This is especially true in countries where teachers are under-trained and under-prepared to do so (such as the US and the UK). The teachers' assessment procedures, on the other hand, were typically unsupported by their instructional aims and required a low degree of intellectual processing. Many teachers were incompetent at evaluating the quality of their assignments (Bol&Strage, 1996).

2.1.2. Knowledge, conceptions, and classroom-based practices in teacher assessment literacy

A few decades ago, Stiggins (1991, p. 3) said that 'the moment has come to encourage assessment literacy for everyone. Assessment literacy has emerged as a hot topic since then, with a growing number of educators looking into how it may be utilized as a standard for teacher education programs in both general education and the teaching of foreign languages (Hamp Lyons, 2016). The ability of stakeholders to use assessment for both learning and

grading is known as assessment literacy (Taylor, 2009). Teachers' ability to organize and develop well-constructed assessment tasks and then use the data gleaned from those tasks to inform pedagogy and learning in a broader socio-cultural context that focuses on teacher assessment literacy (TAL) (cf. Fulcher, 2012). TAL refers to teachers' knowledge, conceptions, and execution of lecture hall assessments in L2 writing contexts to facilitate our conversation.

In lesson planning, there is a significant amount of research that focuses on how instructors are trained to administer classroom assessments. Brown and Bailey (2008) discussed language testing courses. According to this study, assessor training is an essential part of all Hong Kong teacher education programs. In addition, research has been conducted in many international contexts to examine teacher assessment methods. It was discovered by Qian (2014) that English teachers in Hong Kong lacked marking abilities when it came to analyzing student speech in a school assessment. When it comes to doing summative and formative assessments, DeLuca and Klinger (2010) found that Canadian teacher candidates were more familiar with the latter. Vogt and Tsagari (2014) conducted a large-scale European study and found that most teachers had inadequate assessment training and relied on their own on-the-job experiences to evaluate students. In this research, we learn how teachers were trained in assessment literacy and how they evaluated pupils. However, we do not know if secondary school teachers have the competence to deal with standardized testing and classroom-based evaluation, if they are willing to experiment with various assessment practices such as AFL1 and AAL2, or whether they know how to use alternative assessments to enhance their teaching.

2.1.2.1. Knowledge base

TAL's cornerstone is its knowledge base. Students who need to prepare for standardized testing or deliver classroom-based assessments need assessment knowledge. This involves the ability to create, implement, grade, and provide feedback on assessments to enhance student learning. Over the years, several research studies have been conducted on enhancing teacher assessment knowledge through coursework, professional development activities, on-the-job training, and self-study through textbooks. (Harding & Kremmel, 2016). Although more excellent training in assessment is being sought, most teachers are still ill-equipped to conduct classroom assessments with self-assurance and professionalism

(DeLuca & Johnson, 2017). According to a study, assessment course patterns are catching up to practitioners' knowledge needs (Brown & Bailey, 2008).

Further research is being done to see if university coursework can provide teachers in training with current assessment skills (DeLuca, Chavez, Bellara & Cao, 2013). For teachers, this knowledge base has been too conceptual and instructional non-relatable for ordinary classroom assessment methods, regardless of textbook trends and coursework features (Popham, 2009; Yan, Zhang, & Fan 2018). Moreover, the knowledge mentioned above the base is primarily decontextualized, showing that teachers typically learn about relevant assessment knowledge in a cookie-cutter fashion (Leung, 2014).

Due to a lack of assurance and evaluation concepts in L1 and L2, the concept of TAL only has a low profile (Crusan, Plakans & Gebril 2016). According to some academics, it is not enough for language teachers to learn to write revisable written feedback and use that feedback to improve their writing instruction if they are merely given a general assessment course (Lee, 2016). Language teachers in Europe's EFL schools have had to learn about assessment on the job and use educational content as evaluation methods to address the lack of suitable assessment training (Vogt & Tsagari, 2014). Though professional development (a knowledge foundation) was given to both teachers before implementing the focused strategy of written comments and peer assessment in Grade 9 classrooms, the two instructors encountered difficulties.

2.1.2.2. Teacher conceptions

When it comes to teaching, assessment is defined as an internal conceptual basis for how teachers interpret the aims and uses of assessment. Examining teachers' perceptions of the assessment revealed cognitive and emotional aspects. The former represents teacher belief systems, whereas the latter represents teacher emotions when judging student performance.

Cognitive studies show that instructors are better at AoL (Assessment of Learning) than AfL (Assessment for Learning) because they are influenced by the psychometric paradigm of assessment, emphasizing fairness, reliability, and standardization in scoring (Brookhart, 2011; DeLuca & Klinger, 2010). According to these findings, teachers still emphasize standardized testing despite the change from physicochemical to hermeneutic

assessment paradigms. External responsibility and a more comprehensive social norm, such as an exam-driven culture, may blame these teacher mindsets (Lee & Coniam, 2013). To use the phrase "testing as you were tested," teachers' views are often affected by their own experiences in school (Vogt & Tsagari, 2014, p. 391). This idea may influence teachers' evaluations of student writing.

In their minds, instructors tend to see assessment as a bad thing since it is typically associated with the acquisition of specialized knowledge (such as scoring systems), participation in high-stakes decision-making, and overemphasis on summative evaluation. The Education Assessments in Hong Kong presented a professional challenge for English secondary school teachers, according to a study conducted by Qian (2014). Despite training, the teacher participants claimed that they could not assess speaking competency accurately; hence they did not prefer the evaluation. Xu and Liu (2009) found that Betty, a teacher participant, was forced to lower her pupils' participation grades due to the apparent unequal power relationship at play. Because of Betty's superior's directive, she thought it unethical to lower the grades. Shelley, the single instructor participant, was split between policy reforms and actual learning needs, according to Gu (2014). Because she could not risk sacrificing student exam results in favor of experimenting with AfL, Shelley's assessment methods were determined by the content and standards of the public examinations.

2.1.2.3. Assessment practices

TAL's other key component is assessment procedures, which pertain to how creative writing assessments are implemented in instructors' work environments. Teachers' expertise and beliefs are intertwined in these applications. During Lee's (2013) research, educators experienced a shift in identity, from teachers to change agents. Teacher informants were empowered one year after obtaining training in writing teacher innovation, including genre-based teaching, process composition, blogging, and custom rubrics as feedback. Findings from this study revealed the need for teachers to build a firm foundation of knowledge and shift their assessment paradigms to implement alternative assessment approaches (student learning benefits) effectively. Rather than delivering constructive criticism, the instructor in Hamp-Lyons (2006)'s study provided the student researcher, using, with quasi input such as indirect inquiries. However, despite the instructor's continuing pedagogical support, the results showed that she should openly adopt her assessment function when providing

feedback, highlighting Esing's writings' strengths and faults. The instructor was meant to learn well how to effectively manage her roles as a teacher, an assessor, and a language supervisor throughout various stages of the process of writing.

Teachers in New Zealand found that peer evaluation and self-monitoring failed to improve students' writing skills in grades 5 and 6 (Hawe and Dixon, 2014). When Audrey introduced a student-centered approach to writing assessment, she did not adjust her assessment techniques. She rarely communicated her students' tacit assessment criteria to them. Two university professors and six students had varying expectations for written feedback on tasks, according to Hyland (1998). According to her research, students and teachers must engage in a dialogue to determine which sort of feedback is most effective at enhancing textual quality. Teachers need to adapt their attitudes and practices while implementing new evaluation methods to meet their students' requirements and the expectations of their institutions.

2.1.3. Teachers' assessment literacy conceptual framework

Teachers in Hong Kong secondary schools have been studied about their writing assessment knowledge, beliefs, and practices. The teacher participants viewed AoL and AfL as having a basic understanding, whereas AaL lacked. Most of their views on writing exams were positive, believing that they may help students improve their writing. The teachers also tried to implement alternate writing exams but met institutional impediments. In summary, the teachers who took part in the study had a basic understanding of TAL but a need to expand their knowledge of AaL and their understanding of their extra duties as writing assessors. Participants also claimed to be unaware of how assessment data may be used to improve writing instruction.

Teachers in Hong Kong have undergone a two-decade-long process of reforms. However, some teachers remain unprepared to implement alternative assessments or receive adequate training on language assessment, particularly the three complementary facets that can improve students' long-term academic performance. This phenomenon shows that more could have been done to equip teachers of second language writing to be more independent in their assessment of writing and to become more assessment-capable in L2. Three concrete suggestions for TAL development are offered here. First, when teachers are supported in attempting alternative evaluations such as self-reflection in asset allocation activities,

participatory use of shining examples, and application of response dialogues in writing by their principal, they will be more likely to employ these methods. All of these activities require additional financing, instructional assistance, and professional development. Teachers may struggle to fully develop their TAL if they don't have the support, resources, and freedom to do so. As a result of the data, it appears that developing a center of excellence for the exchange of evaluation methodologies is an option. Teachers can establish expert assessment judgments when evaluating writing by engaging in long-term professional conversations and passing on effective assessment techniques to their colleagues. Increasing the professionalism of teachers in writing evaluation is hence likely to aid in TAL's attainment. This study's third finding is that mandating teacher training qualifications, such as BEd and PgDE, to include TAL would have a beneficial spillover effect on student learning and instruction in teacher preparation courses. Students in both courses and training practicums could be assessed on their ability to teach as long as they included TAL as part of their pedagogical expertise. The study's exploratory nature means that its findings may elicit new issues for future research. There are many ways in which L2 writing participants establish and then solidify what comprises the concept of TAL. What and how are pre-service secondary teachers' evaluation literacy skills evaluated in a larger L2 writing context by program administrators? Language assessment training can significantly impact the identity of a writing assessor. Listening and speaking literacy research is expected to benefit from the result of this research, given that TAL is a required but not sufficient prerequisite for the successful use of AfL and AaL in L2 writing. Professionals, educators, parents, students, language evaluators, and authors all have a role in defining, updating, or even negotiating AfL and AaL in writing. Though theoretically significant, this study has certain limitations. Despite using data triangulation, the sample size is minimal, and the self-reported data is susceptible to bias.

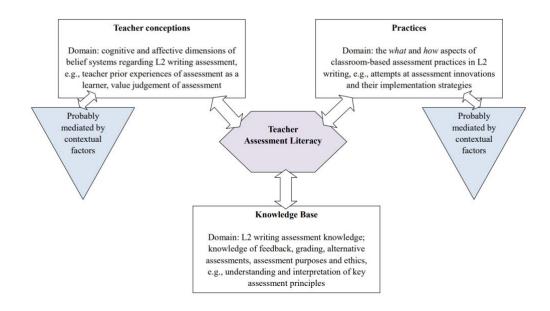Figure 1: Teacher assessment literacy: An adapted conceptual framework (Ricky Lam, 2019)

2.1.4. Teachers' assessment literacy on boosting learners' autonomy

Teachers have a critical role in the education of students. Students' progress is evaluated mainly through teachers responsible for selecting resources, arranging the classrooms, and providing learning opportunities for students. Therefore, teachers' sense of professionalism and assessment literacy is critical in determining the efficacy of any educational setting. On the other hand, instructors and researchers are increasingly focusing on teaching students self-reliant. As a result, this research aims to find out how teachers' professional identities and assessment literacy affect students' sense of agency. The teacher's self-conception is founded on his or her own experiences, opinions, and beliefs about teaching (Ibarra, 1999). "Teacher professional identity" is defined as a teacher's understanding of "teaching-related events and connections that present themselves in supervised practical actions," according to Timotuk and Ugaste (2010).

Bullough (1997) sees the foundation of teacher choices as a teacher's identity. Various sociocultural factors influence how teachers act and behave (Mofrad, 2016). (Abednia, 2012). According to Caihong (2011), teachers' professional identities influence their perceptions of student behavior and improvements, as well as their teaching approaches. According to Ivanovaa and Skara-Mincne (2016), effective teaching is linked to a strong sense of professional identity. Because of this, successful teachers have solid professional

brands. One of the primary goals of teacher training programs is to enhance educators' professional identities, argue van Huizen et al. (2005).

Teaching quality and teacher pedagogical topic knowledge have been linked to teacher professional identity (Veisson & Kabaday, 2018) (Atai & Khazaee, 2014). It has been shown that teachers who have a strong brand and reputation have better relationships with their students and are better able to address their students' needs and solve problems and impart information (Levine, 2006). Two of the essential ideas in language teaching are assessment and the opportunity for students to practice their new skills. Language acquisition appears to require more than just creating acceptable settings and giving accessible input. Learners' strengths and weaknesses can be discovered through assessments. It aids educators in devising, implementing or altering instructional strategies. Similarly, evaluation determines whether or not a teacher is doing a good job (Stiggins, 1999).

As a result, teachers' knowledge of assessment directly impacts the outcomes of their students (Bennett, 2011). It was suggested by Messick (1989) that teachers should have a correct and appropriate understanding of the outcomes of assessments. Assessor literacy was also coined by Stiggins (1991) and is defined as a faculty member's ability to understand and use suitable evaluation methods. "The ability to grasp the many objectives and methods of assessment to find the most appropriate form of test to meet a certain purpose" was described by Ainsworth and Viegut (2006, p. 6). (p. 53). Language assessment literacy was characterized in applied linguistics as the ability "to design, develop, and critically evaluate tests," as well as the ability to "monitor, evaluate, grade, and score assessments based on theoretical knowledge" (Vogt & Tsagari, 2014, p. 377).

Stakeholders in language education and assessment must have proficiency in Language Assessment Literacy (LAL) (Taylor 2009; Yastbaş & Takkaç, 2018). Scarino (2013) argued that a teacher's LAL should be developed. LAL aids educators in self-reflection and self-evaluation of their practices, beliefs, attitudes, and perceptions (Scarino, 2013). The low level of LAL, on the other hand, leads to a flawed assessment of a student (Suah, 2012). This study's third and final variable is the degree of learner autonomy. Taking care of one's education is autonomy (Holec, 1981, p. 3). According to Nunan (2003), Autonomy is a student's ability to keep track of their progress. There are three levels of

autonomy, according to Benson (2001). Monitoring, content, and cognitive processes all fall into these three categories (Benson, 2001).

Autonomous learners, according to Little (2009), work on their projects. In Holec (1981), he says that self-directed learners can create goals, pick materials and methods, regulate their learning, and evaluate their success in the classroom. Ghasedi et al. (2018) found that more autonomous students are more likely to engage in classroom discussions and negotiations. Students that have a higher degree of self-direction are more engaged in their language learning, according to Zarei and Gahremani (2010).

2.1.5. Assessment literacy skills for EFL teachers

Language tests are used in educational and research contexts, yet assessing literacy abilities is required by practitioners to construct linguistic evaluations. (pp. 19–20). Bachman (2000) noticed this mismatch. As a result, Popham (2001) asserted that teachers are not expected to be experts in item composition or capability and organization because their primary role is to teach, not to test. The preparation and grade supervision of summative assessments that are part of standardized curricula in many EFL settings is the responsibility of language teachers (Jin, 2010). In addition, teachers of foreign languages are expected to use the findings of these examinations to improve their teaching and track their pupils' progress. The preparation and application of language assessments necessitate that language teachers have relevant assessment literacy abilities and expertise.

Language and literacy are typically lacking in the career development given to these teachers. Many graduate schools in applied linguistics worldwide do not require students to take a course on language testing because it is not a requirement for master's or doctoral degree students. For instance, in China, though different language instructors at higher institutions are needed to hold a postgraduate degree in applied linguistics or a similar discipline, language assessment courses' fulfillment criteria vary among graduate programs (Wang, 2004). Measurement of educational progress is rarely emphasized in programs that involve language classes (Jin, 2010). It is impossible to adequately train elementary and secondary-level language instructors if the training at the tertiary level is considered poor.

Teachers' frequent involvement in test production and lack of standardized evaluation training has led to a prevalent belief that EFL teachers are incapable of generating

appropriate test items and a concern stated by some language testing academics concerning the quality of test items written by teachers. Teachers' exams are generally of poor quality and provide "little insight into success, development, strengths, and weaknesses", according to Alderson (2005) when it comes to foreign language assessment (p. 4). According to Cai (2013), it can be difficult for language teachers to develop tests since they lack training in item writing and other assessment literacy skills. Coniam (2009) found that most of the multiple-choice (MC) items on a professor EFL test in Hong Kong, China, were of poor item facility.

Even while training language teachers in assessment ideas can help them write better test items, the quality of test items can only be ensured if instructors and test developers adhere to a systematic assessment procedure. All test items must be subjected to a systematic literature and revision process, regardless of the assessment abilities and expertise the item writers have, to ensure their reliability and validity. A valid assumption can be made about the psychometric quality of local primary and secondary EFL examinations because they are often produced in a short period and used just once (Cunningham, 1998). Many EFL professors rely on MC items because of the vast number of exam participants. It is not always easy to develop new ideas for MC products (e.g., Alderson, Clapham, & Wall, 1995; Heaton, 1982; Hughes, 2003). For teachers who have little or no prior information expertise or systematic assessment training, creating plausible irrelevant features for MC items can be extremely time-consuming. Another factor that influences the quality of MC questions when piloting and pre-testing is not implemented: is post-hoc item assessment, which enables teachers to identify problematic things and influences their exam scores uses. The quality and effectiveness of standardized language assessments will be a worry for EFL education in China if these items are not working as they are supposed to function.

2.1.6. Quality of EFL testing

EFL learning and teaching in China, like in most Asian EFL countries, is heavily exam-based. People in China have been learning English as a second language (EFL) since the mid-1990s, and the quantity of EFL courses in China has grown tremendously since then (Cheng, 2008). Standardized examinations are the most effective and fair technique for selecting applicants from varied backgrounds due to restricted resources, i.e. limited options for higher education. Meanwhile, the high stakes of standardized examinations result from

their very particular contexts. In light of China's educational system's emphasis on testing, utilizing standardized test scores in language instruction is vital. It is, nevertheless, based on this concept that standardized language examinations can be used effectively. Operational test quality control is under-researched and under-theorized, but it accounts for most of the daily effort required to run a testing program. Only a few test developers have discussed quality control in testing literature, and Saville (2012) describes quality assurance as "the planning and installation of activities, that over time lead to better being implemented" (p. 399). Quality control techniques, albeit under-researched, make up the bulk of the daily work needed to operate a testing regime (Yan, Thirakunkovit, Kauper, & Ginther, 2016, p. 119). To confirm the accuracy of test scores, good quality control techniques are needed to monitor the production and assessment of test items. Standardized language test scores are primarily accepted in China; however, additional research is needed to examine the quality of EFL testing in China (Cheng, 2008). Even while China has a long history of employing tests for selecting applicants, the country's expertise in language assessment, particularly EFL experiments, is relatively new. The creation of appropriate measures for a significant number of test-takers in the 1990s laid the groundwork for the development of language testing as an autonomous discipline in language studies in China, as Jin (2010) remarked (p. 569).

Yang and Gui (2007) advocated for the establishment of high practice for locally and nationally testing institutions in China in terms of test verification and application. Fan and Jin (2013) concluded that the introduction and validation of state EFL tests in China tended to follow methodical test creation and validation procedures.

Several studies in recent years have examined the validity and reliability of national English examinations such as the College English Test (CET), the Test for English Majors (TEM), and the Graduate School Entrance English Examination (GSEEE) other tertiary-level English assessments (e.g., Cai, 2013). Despite this, the efficacy of regional or institutional Test scores used in primary and secondary school is unknown.

2.1.7. The frequently overlooked secondary EFL tests

Foreign language teachers routinely write local secondary EFL examinations with no experience in language evaluation, typically for edge or final English exams. Quality control methods may not be followed throughout test development. According to an informed source in this study, there appears to be a lack of results in the generation or revision of most test

items, who has worked extensively on regional secondary EFL assessments. Secondary EFL examinations are frequently developed under the guidance of an English research and teaching committee, which is composed primarily of currently employed seniors teaching English and is established by the local school district, city, or province's Board of Education.

A month before the test is scheduled to be administered, the participants in the study group meet to discuss and deliberate on the test's content.

There are not any tests or item criteria here. Each team member is allocated to write things in one or two (sub)sections, and the writing is spread amongst them. The entire team will put together the test when the items have been written. Only those defective items are often eliminated, with the bulk of the others being kept. If all goes according to plan, the test will be ready to operate (Liu, personal communication, July 16, 2014).

In addition, Fulcher (1997) believes even the most seasoned item authors need a comprehensive literature and modification procedure for their written test items. Language testing and applied linguistics literature are scarce in analyzing the validity of language exams designed by EFL teachers, despite their relevance. So, to truly comprehend the value of senior EFL exams and, by extension, the assessment abilities of Chinese secondary EFL teachers, this research assessed the quality of an urban English assessment for 8th graders in northern China.

2.1.8. Literacy testing for language proficiency

According to the findings of López and Bernal (2009), language teachers use a variety of assessment methods. Individuals with expertise in language assessment utilized evaluation to promote teaching and learning, while they used purely those without such experience to achieve academic grades. Teacher assessment and grades were treated as being on the same level by López and Bernal, leading them to conclude that this is a restricted method to be applied to linguistics. In addition, teachers in this study used more formal than formative assessment strategies. While master's degree programs do offer language evaluation courses, few language teachers in Colombia can pursue master's degrees; according to López and Bernal, because of this, experts believe that pre-service bilingual education courses should provide more excellent training in language assessment.

The researchers found that 20 of the 27 undergraduate programs they examined had no language assessment courses. The picture becomes even more complex when the discussion reveals that only two of the 27 programs studied were offered by public universities, compared to five provided by individual universities with assessment courses.

Another study by Arias and Maturana (2005), Frodden and Restrepo and Maturana (2004), and Muoz and Palacio (2012) all revealed similar findings with language evaluation procedures. Also, these findings have been found in various countries, including Chile (Daz, Alarcón, & Ortiz, 2012), China (Cheng, Rogers, & Hu, 2004), and Canada (Volante & Fazio, 2007).

For instructors, López and Bernal (2009) recommend that they enhance their language assessment procedures and establish evaluation that is favorable to teaching and learning in their classrooms, in their conclusion. It is critical that all new teachers take approximately a course in language assessment before they begin teaching, and they should continue to develop their skills through in-service training and conferences to foster a culture of language evaluation in the classroom. For more information, see López & Bernal (2009, p.66).

Teachers are expected to be knowledgeable about all areas of assessment so that they can support their instruction and respond to students', parents', and the community's needs and expectations, state Herrera and Macas (2015) in their paper (p. 303, my emphasis).

According to Herrera and Macau, teachers with an acceptable level of LAL integrate instruction with assessment, criticize large-scale examinations, and construct and choose from an available assessment repertory.

Both Herrera and Macau support López, and Bernal urges school districts to provide more LAL chances for their students so that teachers can concentrate on the whole scope of language evaluation rather than only using examinations as a gauge of progress. The authors then suggest that questionnaires might be utilized to tap into teachers' skills and knowledge in reading comprehension for all experiences. Such an instrument alone is not enough to describe and provide teachers with information on enhancing all. Pre-and-in-service teachers are urged to use LAL in both articles. Assessment for formative—to improve student

learning (Davison & Leung, 2009)—could become vital in language education if appropriately trained language teachers, as these authors argue.

2.1.9. Language assessment generalities in education

It required educators and other participants to have more knowledge and abilities in assessment literacy, even though the emphasis was on teaching effectiveness in the literature.

As a teacher, you can use assessment literacy to monitor, record, evaluate, and evaluate students' progress.

It is also becoming more common knowledge that assessments have ramifications for classroom instruction, student learning, and school curriculum (Brookhart 2011, Popham 2009, 2011) and that teachers should be aware of these implications and take an informed stand against them (Popham, 2009).

In 1990, the American Federation of Teachers, the National Council on Measurement in Education, and the National Education Association Practical Guidance for Effective Teaching in Educational Assessment of Students, included a reference to assessment literacy.

In their opinion, instructors may benefit from these principles if they were made more aware of assessment opportunities both within and outside of the classroom.

Teachers should select, design, and assess good evaluations that benefit learning, teaching, and schools in the first category. Supervisors are expected to learn when evaluations have been misused and effectively communicate results to various stakeholders as part of the second strand.

The term "assessment literacy" was later coined by Stiggins (1995) to describe the knowledge and abilities that educational personnel, such as teachers and staff, should possess concerning assessment.

Also included in assessment literacy is knowledge of internal consistency and threats to it, content validity, fairness, the styling of shuttered test tasks, alternative assessments like portfolios, formative assessment, student preparation for tests, and evaluation of English language learners (Popham 2009).

Assessment literacy, according to Popham, is essential if educators are to understand the influence that examinations, particularly external ones, could have on teaching.

As Brookhart (2011) argues that the above standards are not exhaustive enough for classroom teachers, he believes assessment literacy has more to do with the understanding of how students learn in a specific subject, the correlation between evaluation, curriculum and instruction, design of scoring schemes that are evident for stakeholders, administration of externally-produced tests and feedback to improve learning.

It has been noted that other aspects of assessment literacy include everything from basic statistics for educational measurement, students' motivation, and the use of more than one method for assessment (Popham, 2011; White, 2009; Rudner, 2002).

Teachers' literacy assessments have also been advocated to include technology usage (Rudner & Schafer, 2002). The preceding section established that teachers are expected to have steadily increased levels of assessment literacy knowledge, abilities, and principles.

In some ways, the concept of LAL is like that of assessment literacy, but there are some crucial differences. As a result, this paper aims to address two connected but ongoing disputes in LAL: the necessity of building a foundation of knowledge in this discipline and recognizing that LAL means many things to various people. There will then be a discussion of Scarino's (2013) latest arrival to the definition of LAL, which claims that LAL should include the contexts of instruction by teachers.

2.2. Reflective Teaching

Educators are not exactly pioneers of critical thinking and creative practice ideas. Because they have existed for fifty years (Qing, 2009), to counter the idea of "teacher as a technician," Clarke (1995) argues that the concept of "reflective teaching" was born as a response to the assumption that difficulties linked to teaching may be generalized across diverse contexts and so "did not require on-site interpretation or change" (p. 244). "An over-reliance on technical issue solutions typically leads to frustration and disappointment" (p. 257) because it cannot adapt to each classroom's unique problems. Clarke contends that in this approach, the knowledge of teaching is received rather than knowledge put into practice, and classroom difficulties are predictable and regulated. Schön (in 1987 and 1991) argued

that a teacher goes well beyond the simple solution of problems; instead, teachers attempt to better understand their surroundings by posing problems to students.

Scholars view reflection and reflective activity in various ways. No definition of contemplation or thoughtful teaching can be agreed upon, according to Gunn (2010). (p. 2). A teacher's day-to-day classroom instruction and the institutional arrangements in which he or she works with students are crucial to reflection, according to Bartlett (1990). (p. 204). According to Larrivee (2008), reflection involves acknowledging, articulating, and questioning one's own beliefs. Teaching and teacher education can be improved by reflecting critically on one's own experiences in the classroom, according to Qing (2009), a definition that focuses on instructors' ability to learn from their own mistakes (p. 36). According to Qing, reflective teaching urges EFL teachers to pause, analyze, and inquire about what they are doing. Students are instructed to link theory and practice, analyze both previous and new teaching experiences, and make interpretations about the situations they encounter. On page 36, A reflective teacher is not only concerned with teaching techniques but also with the larger goals of education. As stated by Bartlett (1990), this means that teachers must "move away from the "how-to questions", which have only a practical utility, to the "what to "and "why to "questions." (pp. 205-206).

2.2.1. Forms and levels of reflection

Reflection is defined in the literature in a variety of ways. "Reflection-in-action" and "Reflection-on-action" are two types of reflection discussed by Schön (1991), drawing on his earlier work in 1987. "Both ordinary people and professionals often think in what they're performing, sometimes even while doing it," he remarks. After being surprised, individuals refocus on the action and the implicit knowledge that comes with it" (p. 50).

According to Schön (1991), this type of reflective activity occurs only in the "action-present," which is the period wherein actions still can have a positive impact on a given situation. Schön contends that we do not give our activities much thought when they provide the desired consequences we were hoping for. As a result, "we may respond by reflecting-in-activity," which concentrates on a variety of aspects of action, including "the outcomes of action, the action itself, and the instinctive realizing implicit in action." (p. 56). Reflection-on-action can also occur when it does not immediately link our analysis to current acts; it occurs whenever we reflect on our previous activities to analyze and understand them.

Researchers employ a variety of terminologies to explain the different degrees of reflection that can be achieved.

Van Manen (1977) describes three stages of reflection, with the first level being technical, in which means are more essential than ends. To reach a specific goal, there are a set of rules that must be followed. When it comes to "cultural and individual experience, meaning, perceptions... preconceptions... pre-judgments... for the goal of guiding practical actions," reflection is a two-tiered process (p. 226). Lastly, on the highest level of critical reflection, "worthwhile educational aims in self-determination, community and because of equity, equality and freedom" are sought critical reflections (p. 227).

Four different ways of looking at the topic are being discussed by Boody (2008). First, "the teacher reexamines and ponders over the earliest moments to make some sense of that as well, to learn from it and, presumably, becoming a better teaching assistant" is "reflective retrospection" (p. 500). Secondly, Dewey's introspection is critical thinking. It is an ongoing process in which the teacher becomes conscious of a class problem, analyses the dilemma and analyses it, provides some solutions, tests them, and ultimately enjoys a feeling of enjoyment because of fixing it. This is the third type of reflective thinking: Van Manen's leads to a negative, which comprises "exploring what is educationally worthwhile and providing the circumstances which might allow all individuals to equally participate in the conversation on what is of most worth" (p. 501).

As a last kind of contemplation, Boody points to Schön's "reflection-in-action." Reflection includes a variety of levels and a hierarchical classification, spanning from the simplest type of reflection, which describes classroom behaviors, to the deepest kind of reflection, critical reflection.

Educators and researchers use various terms to describe various levels of contemplation (Alger, 2006; Larrivee, 2008). There are three kinds of reflection, according to Alger (2006), namely descriptive reflection, interactional perception, and critical analysis. In a descriptive reflection, a teacher describes the events in a classroom but does not explain why those acts occurred. Teachers engage in dialogic reflection when they question their views and actions to understand themselves and their students better.

They question what they have always taken for granted at this point. Criticism extends beyond acts and "honors many perspectives" and "expresses the understanding that events

or deeds may be shaped by cultural or political reality," which is the most gratifying of all (Alger, 2006, p. 294). Larrivee (2008) divides into four categories: pre-reflection, surface, educational and critical—all of which are concerned with the problem at hand. Using Larrivee's example, "teachers react to pupils and school events reflexively, without conscious thought of alternate solutions" (p. 342). Teachers focus on "what works rather than any examination of the value of goals as ends in themselves" in mesmerizing, which is the same as technical or descriptive reflection (p. 342). As Larrivee points out, "teachers reflect on educational aims, the theories driving approaches, and the linkages between academic standards and procedures" in pedagogical reflections (p. 343). Finally, teachers use critical reflection to examine both their work and the situational relations in which it occurs.

2.2.2. The importance of teachers' views

As a result of reflecting on their classroom activities, teachers become more aware of what they do in their courses and their steps to cope with various classroom challenges. Reflection (Richards & Lockhart, 1994). "These understandings and beliefs operate as interpretative lenses through which starting instructors make sense of their experience," according to many studies, which stress the need to collect these ideas and assumptions (Alger, 2006, p. 288). "The ideas and processes of thinking which are maintained following the initial incidence of the action," according to Bartlett (1990), when discussing the significance of teachers' pedagogical ideas and the implications on teachers' practice (p. influenced teachers' activities 203).

Donaghue (2003) argues that teachers' attitudes affect whether or not their students adopt new ideas, tactics, and activities. In reality, teachers' ideas about teaching are shown because there will be a considerable gap between input, absorption, and product in teacher professional development, which shows that not all the input instructors are given in instructional strategies will be carried into practice. For various reasons, a few of these concepts are screened out, including the disparity between what students say and what professors believe personally (Donaghue, 2003).

According to Amobi (2003), instructors' views and preconceptions about education influence their teaching style. A strong interdependence between theory and actualized practice characterized education. A person's set of educational values and beliefs influences

their work; work, in turn, influences the values and beliefs of others. This continual activity exemplified the circular nature of the interaction between educational ideology and practice. (page 345).

Teaching, said Richards and Lockhart (1994), is a profession where people can bring their preconceptions about what it means to be a teacher to work. "The foundation to most of the teachers' decision-making, and consequently create what has already been called the environment of teaching," are these individual and varied perspectives and attitudes on education (p. 30).

Teachers' beliefs and values come from various sources, resulting in a wide range of viewpoints and values. As per Richards and Lockhart, teachers' own experiences as language learners may influence their teaching perspectives. Teaching staff bring their personal experiences to the classrooms and "apply their interpretation of those events as paradigms for thinking about an approach to teaching and what instructors do," according to Vazir (2006). (p. 447). The teaching experience of instructors, according to Richards and Lockhart, is the fundamental basis of teachers' beliefs about teaching. Aside from the diverse teaching styles and practices that teachers may have to adhere to, this can impact their philosophy of teaching. In addition, teachers' preferences and personality traits might influence their views on education. Finally, teachers' attitudes about teaching can be influenced by the ideas, techniques, and strategies they implement in their classroom.

Numerous academics have emphasized that teachers' opinions on teaching and education are critical to understanding how teachers make their judgments and that they can be obtained through various methods, such as surveys, interviews, and focus groups. Using the "Repertory Grid Technique," Donaghue (2003) uncovered the ideas of a set of English language instructors from several European countries who had traveled to England to attend two different methodological courses on the theory and practice of teaching English, and the activity's outcomes suggested that it helped [the participants] a lot to reflect of [their views], and about their major approaches to teaching (p.348) As a result, trainers who used it said it "introduced and reinforced the concept of reflection that underpinned the course" (p. 350).

According to Amobi (2003), the "Witcher-Travers Survey of Educational Beliefs," which measures educators' tendencies toward either the teacher or student-centered

instructional belief system, was used to examine the attitudes of 47 first-semester preservice teachers.

A descriptive analysis of the survey results was also included in the qualitative component of the survey. Participants' educational belief systems and their espoused viewpoints were explored in the study's findings and conclusions. After completing the online survey, participants were asked to indicate whether or not the survey categories and analyses genuinely reflected their educational values by agreeing or disagreeing with them. According to the study's findings, there was "high commitment with both the participants' perceptions of questionnaire beliefs and the instructional purpose emphasis of each belief system" (p. 359).

However, as Amobi (2003) points out, the findings of questionnaires designed to elicit instructors' ideas and assumptions about teaching are "only the tip of the iceberg" and are meant to empower new teachers and give them a chance to "have the responsibility of their educational views" (p. 359). For teachers, Larrivee (2000) stresses the importance of examining one's own pedagogical beliefs and assumptions and connecting such views to one's classroom practice. Larrivee points out that "if teachers grasp onto tactics without assessment of what kind of classroom practice will be coherent with their principles, linked with their defined teaching framework, and harmonic with their personalities, they will have only a bag of tricks" (p. 293).

2.2.3. The advantages of reflection

Reflection is the first step toward becoming a more professional person. For a good educator, "the reflection is a trademark," said Bean and Stevens (2002). (p. 205). Teachers can gain insight into their profession through reflection, be critical of it, determine what worked and what did not, and then improve it as a result (Alger, 2006).

"The judgment that is made during pondering on an inspection of a unique experience in light [of] past expertise and experience, generates an enhanced understanding of the implications that will come from the present circumstance," says Amobi (2003) about the benefits of reflection (p. 346). "Reconstruction...is recreating past conceptions and memories to comply with the needs of modern educational settings," says Vazir (2006). According to several academics, reflection is a valuable and productive way for instructors to gain self-

awareness. According to Pollard et al. (2008), knowing oneself is a critical component of reflective teaching. They note that "classroom life exposes these very early" that teachers have "weaknesses and strengths" (p. 104). A "constructive and aim manner" and an "integrated capacity to alter and develop" are two characteristics of reflective teaching, they say (p. 104).

As Larrivee (2008) points out, critical reflection requires an in-depth analysis of oneself, which includes a critical examination of one's ideas, preconceptions, expectations, and values. Reflection can also assist teachers in breaking free of the monotony of the classroom, which can be a source of anxiety for some.

In Qing (2009), he argues that teaching experience helps teachers create their instructional methods of dealing with the classroom's daily routines. "although a teacher's teaching style can help them cope with many of the day-to-day challenges of education, there seems to be a risk that it can impede a teacher's professional advancement" (p. 35). Despite gaining from their experience at school and tying them to theory, teachers' practices would never be called professional, according to Qing's advice (p. 39). According to Gunn (2010), teachers might avoid "sliding into an attitude of regular, repetitious "one-tier board all teaching" by maintaining a reflective practice. (p. 208).

2.2.4. Reflective teaching adaptation

An intervention formula is not enough to help teachers become critical thinkers, according to Larrivee (2000). It would be best if you experienced it firsthand (p. 306). Teachers who engage in reflective practice must critique and reflect on it. There are several considerations for potential reflective practitioners to keep in mind. There are two approaches for students to become thoughtful about their behaviors in the classroom: "Questioning and problem-solving" (p. 57). Critical reflective teaching requires a teacher's self-awareness and self-control, as well as an awareness of their instruction's "historical, social, and cultural context" (Bartlett, 1990, p. 205).

According to Bartlett, a teacher's ability to deal with current events and structures — and not take them for granted — is one goal of being a critically reflective teacher. (p. 205). Becoming a critically reflective teacher involves realizing that we are both the makers and producers of our history. The practical meaning of this is that systematic and social

techniques of the investigation will be used for us to see the variables that inhibit change and hence improve everyday teaching. (p. 206).

Teachers who want to become reflective practitioners must work on their personal development. If instructors are going to be reflective, they need to show open-mindedness, accountability, and wholeheartedness, according to Pollard et al. (2008). Because any inquiry that knowingly depends on partial evidence only weakens itself, progressiveness is a necessary trait for rigorous thinking" (p. 20). Reflective teachers need to be "ready to reflect upon [their] own assumptions, prejudices, and philosophies, as well as those of others," they write (p. 20). Pollard et al. clarity that while accountability refers to evaluating the repercussions of each instructional action, wholeheartedness "refers basically to the way within which consideration takes place" regarding the other expected dispositions of a practitioner (p. 20).

According to Stanley (1998), teachers may find it challenging to maintain their devotion to reflection if it reveals negative aspects of their work. If you look deeper into your teaching, you are more likely to identify concerns of discrimination or favoritism toward certain trainees, learning styles, or ideas of teaching and education. " (p. 587). Because of this, Stanley says that "putting oneself and one's work under scrutiny requires a reasonable amount of ego development" (p. 586).

2.2.5. Reflecting teaching cycle

Teachers who want to engage in reflective practice must do so in a nonlinear cycle. So, teachers may need to go through some portions of the cycle more than once or skip others when reflecting on their practice, as noted by Bartlett (1990). 'Mapping, informing, contesting, appraising, and acting,' says Bartlett, are all part of this cycle. A map describes gathering data through various means, such as journaling and recording audio or video, whereas a piece of information to pursue fundamental principles in education.

To put it another way, for instructors, "[informing] begins the initial steps toward acknowledging [their] ambivalence about [they're] commonly accepted and most beautifully held views about education and its growth of organization" (p. 211). To "look for glaring inconsistencies in what instructors do and how they believe" is to "contest" the preconceptions about teaching (p. 212). Teachers' "new understanding," as described by

Bartlett (p. 213), must be "compatible with [their] new appraisal" (p. 213) before they can "act" on their new teaching understanding.

2.2.6. Reflective teaching assessment

We might consider early classroom studies of language teaching in the 1970s an example of a procedure approach to education. It was determined how effective a particular teaching method was based on the teachers' teaching and the student's achievements. If the processes were determined to be successful, they would be categorized as supportive ones that could be used by other instructors (Borg, 2006, pp. 5-6). However, publishing a paper by the Institute of education in 1975 signifies the beginning of critical reflective research that focuses on instructors' thoughts. "Thus, the report recommended that teachers reflect on the connection between their thoughts and their actions" (quoted in Borg, 2006, p. 7). As a result, academics considered and investigated instructors' ideas as a factor in their teaching behavior.

Literature also shows that critical thinking and reflective teaching have been studied together. Reflective teaching and critical thinking, according to some researchers, have their roots in critical thinking (Cornford, 2002; Hillier, 2005). "The technique through which teachers take part in aspects of critical thinking, such as rigorous deliberation and analysis, making choices, and deciding on a course of action related with teaching," says Van Manen (1991; cited in El-Dib, 2007, p. 7).

A study by Gobena (2016) explored the attitudes of principals, supervisors, and mentees to the research process as critical reflection. The study's findings revealed a link between lecturers' attitudes regarding action research as a reflective activity and a lack of creative approaches, instructional resources, and faculty commitment. For this reason, school leaders and supervisors in the system's implementation units must adhere to high standards of conduct in these areas to meet the country's professional needs while also ensuring that Ethiopian students receive an education that meets their needs.

2.2.7. Reflective teaching practices on students' academic achievement

Students' academic accomplishment is an essential measurable measure for assessing the consequences of schooling, and it is rightly considered a primary goal in the world of

education. (Rastogi, 2012). To achieve any educational system's goals, pupils must be able to succeed academically, and every educator is continuously looking for causes for academic performance variation. Since teachers are ultimately responsible for putting educational policies into action and exert a significant influence on pupils, they are at the heart of each education reform effort.

According to the National Education Law of 2014, "Because teaching practices are the most critical determinant in students' desired learning outcomes, the design of the in professional development should stress quality and applicability. (The Education Policy Strategy and action plan, 2016-2021). Student achievement was boosted as a result of a variety of teacher-related factors.

Teachers' cognitive abilities, topic knowledge, teaching and learning expertise, professional development, and classroom behavior are all part of this package (Fong-Yee& Normore, 2013). It has been shown that teacher quality has the most significant impact on student ability when contrasted to other variables such as class size, subject, students' socioeconomic background, and the student's home environment (Rivkin, Hanushek, and Kain, 1998).

Teachers, according to King and Newman (2000), should be given the highest priority when it comes to developing students' abilities, knowledge, and attitudes since they are the ones who have the most direct and significant interaction and influence on students' learning processes and environments. As a result, we must place a higher value on teacher development to boost student achievement.

To become an effective teacher, one must have diverse abilities and attitudes that can only be developed through formal training and on-the-job experience. Teachers can improve their craft by gaining more experience, but this does not guarantee that they will become better teachers. Classroom and school tasks like paperwork and other "regular demands of teaching" can prevent teachers from "achieving a better level of understanding of how they educate" (Richards, 1995).

Educators can avoid this by developing a habit of self-reflection and using their metacognition skills in their professional growth. As a teacher, one must be a skilled observer, an effective communicator, and an expert in the subject area, and improve their

professional skills based on the mix of scientific and academic understanding, personal experiences, and preferences (Cephe, 2009).

As a result, teachers are urged to know the scientific understanding of teaching and implement the practices that are regarded as most beneficial for their students. To accomplish this, teachers must be able to analyze problems in teaching and learning from several angles and come to rational conclusions, rather than relying on guesswork, haphazard ideas, and commonly accepted answers.

The recent educational reforms place a high value on teacher preparation. Historically, the school curriculum in Myanmar has focused on imparting factual knowledge to students. The use of "closed questions, cued stimulates the activity, and choral responses" (Hardman, Aung & Myint, 2012, p. 9) in teacher education in Myanmar may harm students' analytical thinking abilities."TCSF was established in 2016 to address these difficulties and reforms to address them as part of the school.

Professional awareness and expertise, professional development the skills, personal practice and dispositions, and personal growth are four pillars of the framework for the positive impact on a teacher's productivity (Aye Aye Myint, 2016).

Currently available programs and training for the professional development of in-service teachers are woefully inadequate for implementing this paradigm since they focus primarily on teaching the content of textbooks rather than addressing instructional issues. Lacking continuity, instructor involvement, and practicality is this strategy.

Teacher training in Myanmar is a common practice in countries with little resources, such as Burma. Consequently, the knowledge distributed is not entirely extra to all educators (Hardman, Aung & Myint, 2012), which leads to a deficiency in teachers' capacity and subsequently to poor student learning. The learning programs are conducted in short periods with randomly selected school teachers who must teach their coworkers back in their schools.

Reflective practice should be a part of teacher education to apply theoretical knowledge in the classroom more effectively (Hardman, Aung & Myint, 2012). In other words, the instructors in Myanmar classrooms need to be self-reflective and committed to their professional growth. They are conscious and questioning their assumptions and values

and are aware of and attentive to the circumstances in which they teach which they work (Zeichner & Liston, 2013).

Effective teachers help students learn more effectively. Reflective teaching is critical to the development of novice instructors, so according to Kheirzadeh & Sistani (2018). There is a correlation between pupils taught by competent teachers and their academic performance. Even while it is often assumed that reflective teaching helps pupils do better in school, no one has studied this claim.

2.3. Thesis Reflective Teaching and Assessment Literacy

2.3.1. EFL teachers' assessment literacy and their reflective teaching

The complexity of teaching calls for professionals with the appropriate training. Education systems are judged on the quality of their teachers. As Campbell (2000) points out, the teacher is a critical component of the language learning process. Assessment methods may be one of the most crucial components of the teaching process, as numerous research has shown that assessment is used to boost learning in the classroom (Earl& Katz, 2006).

As a result, teachers' work includes regularly assessing their pupils' development (Brumfit & Johnson, 1979). Even while English teachers need to be aware of alternative techniques of testing students, many of them still use the old standbys. As a result, an increasing number of teacher and administrator professional development programs have focused on assessment literacy, and many scholars and organizations have identified specific curriculum areas in which instructors must enhance their assessment skills.

John Dewey first proposed retrospective thinking in 1933, and it has since been extensively investigated. Dewey outlined the concept of reflective thought, defining it as the active consideration of any idea in light of previous knowledge and future aims. "Wherein educators and teaching assistants gather information about learning, monitor their etiquette, beliefs, theories, and teaching methods, and use the details achieved as a founding for critical analysis about teaching," said Richards and Lockhart (1996), are examples of a reflective approach to teaching (p.1).

Based on this statement, instructors can learn more about teaching and how to teach better by thinking critically about what they have learned from their teaching experiences. It is the job of a reflective teacher to consider the genesis, principles, and outcomes of their students' classroom work. An improved understanding of the practice and steps to improve practice arise from these focused discussions (Rose, 2007). "Professionals are confronted with a circumstance that they regard as distinct or bearing a constituent of surprise," says Griffiths (2000).

Rather than relying solely on their knowledge and experience, experts use their own experiences to reinterpret the problem and come up with new ideas" (p. 542). However, Schon1987 argues that "teachers' conscientious concerns and retrospective assessment of their activities to obtain knowledge from experience" constitutes "reflection upon action" (cited in Leitch and Day, 2000).

2.3.2. Review of reflective teaching and assessment literacy literature

There has been much discussion on the importance of enhancing teachers' evaluation literacy (Deluca & Klinger, 2010; Popham, 2009a; Volante & Fazio, 2007). Sadly, many teachers have not been trained in the evaluation; thus, they have a limited understanding (Popham, 2004, 2009a). According to the National Council of Teachers of Mathematics, a significant issue in education today is the "abysmally low" level of assessment literacy amongst teachers (Popham, 2010, p. 175). In many studies on pre-service literacy teacher education, the necessity of teacher reflection on the process was also studied (Odo, 2015). Reflection can take on a wide variety of shapes and forms, as evidenced by this study. Several studies have found that teachers who engage in self-reflection are more equipped to deal with the issues they face in the classroom (Brookfield, 1995; Larrivee, 2000) and can better integrate their expanding knowledge into a cohesive framework (Odo, 2015). According to Pollard (2002), reflective teachers will be aware of the complexities of assessment and student performance and the ways wherein evaluation can help students learn.

EFL teachers' evaluation of literacy and reflection teaching has become increasingly important in recent years, and this study intended to examine the relationship between the two. The study results showed that EFL teachers' literacy evaluation and reflective teaching are positively correlated. This suggests that the more instructors analyze their students'

literacy, the more they reflect on their teaching practices. Increased evaluation literacy also improves the quality of instructors' self-reflection on how well they are teaching. Teachers' reflective teaching has the most link with their performance, while teachers' reflective teaching has the weakest correlation with teachers' disposition toward assessment.

Finally, theoretical model modeling (SEM) was used to see if instructors' assessment literacy predicted their reflective teaching. The suggested model reveals that EFL teachers' assessment literacy predicts their reflective teaching. The findings of Yazdani, Amirian & Hedayati (2015), which found a favorable correlation between critical reflection and EFL teachers' judgment of students' achievement, were consistent with the results of this study.

As Cole (1997) and Coyle (2002) have argued, reflective teachers analyze, discuss, evaluate and change their practice, eventually leading to student progress. The findings confirm this. In the most concrete terms, this means that many pre-service programs and teacher education programs do not currently presume that their teacher candidates will graduate with sufficient assessment literacy to assess pupils properly (Campbell, Murphey &Holt, 2002).

Consequently, instructors must be equipped with the essential knowledge and abilities for evaluation as part of all educational curricula. Before beginning their careers as teachers, pre-service students and teachers alike should be exposed to a variety of evaluation methodologies as part of their curriculum and career development opportunities (Davidheiser, 2013).

3. METHODOLOGY

Many believe that research is the same. If you believe that neither scientific has a method, then the results of any scientific investigation are equally valuable, regardless of how they were achieved. Throughout Descartes' "Discussion and debate on the method," he emphasizes the importance of the method in persuading scientific truth.

We have always viewed methods and techniques used to identify and prevent slips as part of the method. It is impossible to achieve the objectives of research or empirical reasoning unless the proper approach is used. For example, research is a proven method, not a subject of investigation. To ensure the validity of a researcher's findings, they should know that the method they use has a significant impact.

In scientific writing, the term "research technique" connotes various meanings. The methodology is a system of proven, dependable, and methodical rules, tools, and procedures in analyzing data to identify unknowns, solve issues, and uncover facts. Several tools and authorities are used in scientific research to collect and analyze data to reach a specific purpose: to find reality. Another facet of this science's applicability is comprehensive management with basic quantitative methods to examine the validity of ideas.

Today, management researchers use statistical approaches to test many of their assumptions and examine the accuracy of the correlations between their variables. Using statistical tools ensures the confirmability of all interviews and questionnaire surveys. Standards and procedures are essential in today's volatile environment, and their utilization is inevitable.

Many strategies and questionnaires were searched to get information and analyze data in this dissertation chapter, but they all revolve around the study of a statistical population and how to select a sampling method and sample size. Other concerns discussed in this chapter include confirmability (Cronbach's alpha), data gathering instruments (questionnaires), and software employed.

To sum up, this chapter contains the overall design of the study, participants, research context, data collection instruments, data collection procedures, and discussion.

3.1. The Overall Design of the Study

A researcher's goal in deciding on a method is to decide which methods are most appropriate for a particular topic and which methods the researcher should use to provide accurate and timely responses to their research questions. It can generally split behavioral sciences research methodologies into two categories based on the research purpose and the method used to collect data. Depending on the research goal, fundamental, applied, and developmental studies are all subcategories of scientific research. When it comes to basic research, the primary goal is to test ideas to analyze and explain the correlations between phenomena and add to a particular field of study's understanding. Theories are tested and confirmed, either changed or rejected in this type of study. To get practical knowledge in a particular sector, researchers conduct "applied research". Applied research aims to put into practice scientific concepts and viewpoints developed. In this type of study, the goal is to establish and determine the best model, plan, or program for a particular situation. The primary goal of developmental research is to generate or provide a model of a program or plan that first explains a specific nuclear scenario before the model, plan, or specialized program is built based on the study findings.

The researcher has intended to conduct a quantitative research method to find out whether there is a relationship between EFL teachers' language assessment literacy and their reflective teaching. For this purpose, the sample that is chosen by the researcher includes the EFL teachers who are working at Baskent and Gazi University in Ankara in Turkey. The size of this research focused on aiming at approximately 110 EFL teachers. The identities of these teachers will be kept anonymous owing to privacy reasons.

The researcher aims to implement the questionnaires for exploring whether there is a relationship between EFL teachers' language assessment literacy and their reflective teaching. According to Denzin and Lincoln (2005), questionnaires have been an essential way to measure something because they keep the privacy rights of the respondents. Questionnaires, on the other hand, are a poor method of measuring when researchers are working with larger groups. (Davies, 2007). Sometimes problems might occur if the respondents don't answer all the questions listed, but there is nothing that the researcher can do in this case. Due to COVID-19 pandemic conditions, the researcher had to conduct an

online questionnaire for this study, which reduced the number of participants. Online data was gathered using Google Docs.

The data were analyzed using the IBM SPSS Statistics 23 package application. When analyzing the study's data, categorical variables were given frequencies (number, percentage), and numerical variables were given descriptive statistics (mean, standard deviation, minimum, maximum). Cronbach's alpha was used to calculate the questionnaire's internal consistency reliability.

3.2. Participants

The main respondents and participants of this research are university lecturers from Başkent University and Gazi University. These lecturers are the formal teachers of the school of foreign languages. The statistical population of the present study included 110 people. In the study, there are 31 male teachers (28,2%) and 79 female teachers (%71,8). The biggest part of teachers with 39,3% (42 people) graduated with a bachelor's degree, 29,9% (32 people) graduated with a master's degree, 8,4% (9 people) graduated with Ph.D. and last part 22,4% (24 people) graduated from others. Three participants have not responded about his/her area. The 46,4% of teachers' department is TEFL, the 18,2% of teachers' area is English Translation, and the rest of the group is 35,5% from other departments. Thirty-five participants work at a state university; seventy-five participants work in a private university. And last demographic question was about years of experience, the first group 1-5 years is 19,1%, the second group 5-10 years is 33,6% and the last group 11-more is 47,3%.

3.3. Research Context

The research was carried out with the help of lecturers from Başkent University in Bağlıca/Ankara and Gazi University in Gölbaşı/Ankara. These two institutions in the capital city of Turkey are well-known and famous universities. However, because to the global COVID-19 pandemic, the study was conducted entirely online. The quantitative data was collected using Google forms by the researcher.

3.4. Data Collection Instruments

This was a quantitative research project, and the data was collected using two different instruments. These questionnaires have been used in the past by many researchers and it has

been proven that they are strong tools when the researchers need to access to opinions of a wider community (Patton,2002). Prior to using these two instruments, the researcher asked the respondents six personal questions in order to learn about their ages, genders, grade levels, areas of their degree, type of school, and years of teaching experience.

In the first questionnaire, Akbari, Bahzadpour, and Dadvand have developed an English language teaching reflection inventory that teachers may use to monitor their progress. Participants' reflective habits are assessed in five subscales by completing a 29-item test. For this purpose, a six-component model of the second language (L2) teacher reflection, encompassing practical, cognitive, metacognitive, affective, critical, and moral reflection, was developed using experts' opinions and a comprehensive review of the related literature.

The following six major components of reflection emerged from the questionnaire analysis stage and would be assessed and validated in the study's succeeding phases:

a) Practical element: This component includes those topics that speak to the equipment and the real process of introspection. Journal writing, lesson reports, surveys, questionnaires, audio and video recordings, observation, action research, teaching portfolios, group discussions, and analysis of crucial episodes are some of the several instruments and methods for reflective practice (Farrell, 2004; Murphy, 2001; Richards and Farrell, 2005; Richards and Lockhart, 1994). However, conducting action research has been defined in this study as belonging to the "cognitive part" of reflective practice.

b) Cognitive element: This element is interested in teachers' efforts to further their professional growth. The behaviors included in this area include reading the professional literature, attending conferences and seminars relating to one's field of study, and carrying out small-scale classroom research projects (action research) (Farrell, 2004; Richards and Farrell, 2005).

c) Learner element (affective): This component includes topics related to a teacher's reflection on his or her students, how they are learning, and how students react or behave emotionally in class. This tendency "emphasizes contemplation about pupils, their cultural and linguistic backgrounds, thinking and under- grasp, their interests, and their developmental appropriateness for particular tasks," according to Zeichner and Liston

(1996). (p. 57). This component focuses on teachers' evaluations of their pupils' emotional reactions in the classroom (Hillier, 2005; Pacheco, 2005; Pollard et al., 2006; Richards and Farrell, 2005; Richards and Lockhart, 1994).

d) Meta-cognitive element: This component focuses on instructors' self-reflections on their own personality, beliefs, and practice as well as their own emotional makeup and practice definition (Hillier, 2005; Pollard et al., 2006; Richards and Lockhart, 1994; Stanley, 1998; Zeichner and Liston, 1996). Teachers' personalities, and more especially their psychological makeup, can affect their propensity to engage in reflection and their response to the self-image that results from reflection (Akbari, 2007).

e) Critical element: This component comprises articles that discuss and reflect on the socio-political dimensions of teaching. Items in this category deal with educators debating the political implications of their work, introducing issues of racism, gender, and social class, and investigating strategies to empower students (Bartlett, 1997; Day, 1993; Jay and Johnson, 2002; Zeichner and Liston, 1996).

f) Moral element: Items featured in this verify the moral makeup of the teachers. Valli's (1990) three strands of thought, which take into account the idea of morality, are cited by Hansen (1998). The "deliberative approach" "urges teachers to think critically about their aims and how to defend them from a moral point of view," according to the statement (Hansen, 1998, p. 644). The second strategy, referred to as the "relational approach," "draws on feminist theory and moral philosophy, which emphasizes the moral life around concerns of human character and how people see and treat other people" (p. 645). Goodman (1986), Apple (1979), and Giroux and McLaren (1986) also emphasize the third strategy, referred known as the "critical approach," which Hansen (1998) claims is quite similar to critical reflection (Cited in Hansen, 1998); Since the fifth element (above) dealt exclusively with this feature, items relating to the third strategy were left out of this list.

This subscale, cognitive analysis (six questions), measures teachers' conscious efforts to engage in professional development, such as attending meetings or reading professional journals. The third subscale of the Affective Reflection comprises efforts to obtain more profound knowledge of the students' experiences and elicit their thoughts on instructional activities (three items). Among the seven items in the fourth subscale, metacognitive reflection focuses on instructors' subjective ideas about the nature of teaching, as well as

43

critical evaluation of their performance and judgments of their own teachers' behavior. To sum it all up, the seventh and final subscale, critical reflection, assesses teachers' understanding of and ability to apply socio-political aspects to their classroom practices. The assessment should indeed be answered on a 5-point Likert scale, ranging from (1) "never" to (5) "always," according to the instructions.

The second questionnaire was taken from Vogt and Tsagari's (2014) Teachers' Questionnaire, which was created based on Hasselgreen, Carlsen, and Helness' (2004) Questionnaire. That is, the test included three sections, but a fourth, "Knowledge of Testing and Assessment," was introduced on the basis of Brown & Abeywickrama's (2010) book since it contained items that would be crucial to preservice language teachers' LAL. Additionally, the original Likert scale allowed teachers to provide responses in one of three ways, whereas the modified version required participants to provide responses in one of four categories:None, Little (1-2 days), Sufficient, Advanced.

Overall, the second half of the questionnaire had 4 sections and 112 items after these modifications and revisions. Similar to the original instrument, each of these sections had both Part A and Part B. Participants were asked to respond to the training they had already received in the given item in Part A, while they were asked to respond to the training they believed they still needed in the same given item in Part B. Alternatively put, the items provided in Part A and Part B of each questionnaire part were identical; to account for the training acquired in Part A and the training required in Part B, these questions were posed and required to be answered independently. Additionally, Part A used a 4-point Likert-type scale while Part B used a 3-point Likert-type scale.

36 items made up the first component, which looked at the training received and the need for more training in the *classroom-focused LTA* domain (18 items in Part A; 18 items in Part B). The participants had to respond to questions in Part A regarding the instruction they had received. In other words, participants were asked to score the instruction they believed they had received by selecting from a 4-point Likert-type scale (None; Little (1-2 days); Sufficient; Advanced) in this part. The participants had to respond to the identical questions in Part B on the necessary training. They were essentially asked to rate the training they believe they need for the provided item by choosing from a 3-point Likert-type scale (None; Yes, basic training; Yes, more advanced training). The other 12 items in both parts

were added after carefully reading the books in the field (Brown & Abeywickrama, 2010; Hughes, 2003) to gain a thorough understanding of the extent to which preservice language teachers' received and need training in this domain. Six of these items in both parts were taken from Vogt and Tsagari's (2014) study as stated above.

The researcher prepared the second portion, which examined the training obtained and the need for more training in *Knowledge of Testing and Assessment* domain, based on the book by Brown & Abeywickrama (2010) because these items were seen as crucial and required parts of LAL. This component had 18 questions (9 in Part A and 9 in Part B), and in Part A, preservice language instructors were asked to describe the training they believed they had received in order to answer the question. by choosing from the same 4-point Likert-type scale ranging from None to Advanced. As for Part B, they were prompted to identify the training they believed they required for the objective by choosing from a 3-point Likert-type scale ranging from None to Yes, more advanced training.

The third segment, which had 14 questions (7 in Part A and 7 in Part B), was designed to elicit information regarding the training participants had received and the need for additional training in the domain of testing purposes. The researcher included the remaining 3 items in both parts on the grounds that they were also related to the types of tests as stated by Brown and Abeywickrama (2010) and thus related to the *Purposes of Testing*. Four of these items in both parts were also taken from the study of Vogt and Tsagari (2014). The participants were asked to respond to the provided question by selecting from a list of the training they believed they had received and needed both the 4-point and the 3-point Likert-type scale separately.

Preservice language teachers were asked to describe the training they had received and the need for more training in the *Content and Concepts of the LTA* domain in the last part by selecting independently from the 4-point and 3-point Likert-type scales. There were 44 items in it (22 items in Part A; 22 items in Part B). Two of the items in each part were taken directly from the Vogt and Tsagari (2014) questionnaire; the other five were modified and some changes were made, such as splitting some items into two to make them more specific and rephrasing some items, in order to better understand the participants' LAL and training needs in this area. To ensure that this section included the fundamentals of LTA (e.g*., different test items/task types to test reading in English, Testing pronunciation in English...*), the

45

researcher created the remaining 15 items for the two parts after carefully reviewing the books by Brown & Abeywickrama (2010) and Hughes (2003).

3.5. Data Collection Procedures

Data analysis is a multi-stage process in which data got through collection tools are summarised, coded, categorized, and finally processed to provide a variety of analyses and relationships between these data to test hypotheses.

In these processes, data is refined both conceptually and empirically. Various statistical techniques play an essential role in inferences and generalizations. Analysis processes are unique according to the type of research, research method, nature of hypotheses, tools used to collect information, etc. (Khaki, 2011).

Before the questionnaires have been implemented for the EFL teachers, the required permission is obtained from university authorities. The participants have been informed about the purpose of the study before the questionnaire has been implemented and participation in the study was voluntary. The questionnaires were delivered to the teachers via an online link.

The research technique comprises a series of tasks to reach a goal and discover natural and authentic laws. Therefore, research methodologies are instruments for achieving reality (Delavar, 2004). As part of our fieldwork, we asked our participants to complete a questionnaire to gather data about their experiences. Based on these ideas, the subject literature and the analysis methods of EFL teachers have been utilized to establish research signals and how these variables are used to identify how these variables relate to each other. The EFL teachers were given the questionnaire to test its validity and reliability, and they found it to be both accurate and reliable. After the questionnaire was completed and collected, extraction, categorization, and analysis were carried out. When a specific issue or topic is being studied, such as the classroom environment or its impact on EFL teachers, applied research is the type of research that aims to uncover and implement helpful and successful practices. In terms of research methodology, this study may be classified as a descriptive survey (field) study because it was used to analyze the domain's behavior and fill out a questionnaire.

4. RESULTS AND DISCUSSIONS

This part of the study presents the results of the data analysis to examine the relationship between reflective teaching and language assessment literacy of EFL teachers. Firstly, demographic data are shown in detail. Secondly, the findings of the research questions are given. The findings and achievements of the project are extensively reviewed, discussed, and analysed. In this chapter, for all the studied objectives, the obtained results are presented in order, and their statistical test results are determined. Analysis of this research's text or contents consists of data analysis and potential cause and effects. In this section, all the results obtained are explained.

4.1. Descriptive Statistics

The statistical population includes English language teachers,lecturers or instructors. In addition, considering that the number of people in the study population was 110, according to Cochran's formula, the number of samples, as seen below, equals 100 people. Therefore, 100 questionnaires were distributed among the EFL teachers and collected after completion. Table 4.1 briefly describes the demographic of the research sample:

Table 4.1. Distribution by Demographic Features

Demographic description		Number	Frequency
Education Level	Bachelor's degree	39	39%
	Master's degree	29	29%
	Doctorate	8	8%
	Others	24	24%
Gender	Man	29	29%
	Female	71	71%
Age	21-30	29	29%
	31-40	54	54%
	41-50	14	14%
	51-more	3	3%
Seniority (Years of teaching experience)	1-5 years	19	19%
	5-10 years	33	33%
	11-more years	48	48%

When Table 4.1. is examined, the majority of teachers 39% (39 people) graduated with a bachelor's degree, 29% (29 people) graduated with a master's degree, 8% (8 people) graduated with Ph.D. and last part 24% (24 people) graduated from others. In addition, 71% of the people participating in the study are female and 29% are male. Furthermore, 29% of participants are between 21-30 age, 54% of participants are between 31-40 age, 14% of participants are between 41-50 age and 3% of participants are older than 51 age. Last demographic question was about years of experience, the first group 1-5 years is 19%, the second group 5-10 years is 33% and the last group 11-more is 48%.

4.2. Analysis of Research Data

Here, one should try to present only the contents in the form of tables and diagrams that help to understand it more and are in line with the objectives of the dissertation. A result is avoided in graphs and tables, although it is necessary to use previous research to make the data more accurate and enriching. In order to describe the data, the mean and standard deviation of the research data were used. A summary of the descriptive statistics related to the research variables among the individuals of the study population is examined and is given in the following tables (a code is provided for the questions of each of the variables).

4.2.1. Principal component analysis

Criterion deviation: is a concept that determines the scatter of data in a set and is, therefore, one of the most important statistical scales in the field of descriptive statistics.

Skewness: Skewness is actually a measure of the degree of symmetry of the distribution function. For a perfectly symmetric distribution, the skewness is zero; for an asymmetric distribution with skewness towards higher values, the skewness is positive; for an asymmetric distribution with skewness towards smaller values, the skewness value is negative.

Weighted average: The sum of the means of the sub-indices divided by the number of sub-indices.

Table 4.2 Descriptive Information on the Status of the Reflective Teaching Variable

Skewness	Standard deviation	Average	Maximum	Minimum	Number of data	Research criteria	Component
-.164	1.227	3.22	5	1	100	A1	
-.431	1.143	3.63	5	1	100	A2	
-.026	1.185	2.90	5	1	100	A3	**Practical**
-.516	1.164	3.59	5	1	100	A4	
-.001	1.225	3.07	5	1	100	A5	
.053	1.178	2.84	5	1	100	A6	
-.356	1.123	3.52	5	1	100	A7	
-.219	1.095	3.35	5	1	100	A8	**Cognitive**
.013	1.313	2.82	5	1	100	A9	
-.350	1.062	3.38	5	1	100	A10	
-.112	1.105	3.01	5	1	100	A11	
-.296	1.092	3.28	5	1	100	A12	
-.330	.980	3.78	5	1	100	A13	
-.577	1.090	3.77	5	1	100	A14	**Learner**
-.392	1.105	3.51	5	1	100	A15	
-.686	1.155	3.67	5	1	100	A16	
-.727	1.034	3.89	5	1	100	A17	
-.814	1.020	3.97	5	1	100	A18	
-.752	1.040	3.78	5	1	100	A19	
-.695	1.040	3.99	5	1	100	A20	**Meta-Cognitive**
-.643	1.101	3.86	5	1	100	A21	
-.523	1.092	3.60	5	1	100	A22	
-.323	1.218	3.25	5	1	100	A23	
-.375	1.085	3.44	5	1	100	A24	**Critical**
-.044	1.118	2.89	5	1	100	A25	
.466	1.371	2.40	5	1	100	A26	
-.578	1.050	3.74	5	1	100	A27	**Moral**
-.345	1.198	3.28	5	1	100	A28	
-.354	1.123	3.45	5	1	100	A29	

As can be seen above, most of the Reflective Teaching variable questions have an average above 3. Moreover, the data set's standard deviation is close to zero, indicating that

the data are close to the mean and have little scatter. According to the table, skewness coefficients are in the range (2+ and 2), and the probability distribution Table 4.2. shows this data as normal.

Table 4.3 Descriptive information about the status of the Language Assessment Literacy variable

Skewness	Standard deviation	Average	Maximum	Minimum	Number of data	Research criteria	Question group	Component
-.262	.946	2.71	4	1	100	B1		
-.203	.890	2.58	4	1	100	B2		
-.393	.880	2.71	4	1	100	B3		
-.264	1.010	2.70	4	1	100	B4		
.091	.904	2.47	4	1	100	B5	Please	
-.207	.940	2.62	4	1	100	B6	specify the	
-.166	.998	2.57	4	1	100	B7	amount of	
-.388	.974	2.80	4	1	100	B8	training	
-.407	.939	2.87	4	1	100	B9	you think	
-.407	.991	2.74	4	1	100	B10	you have	
-.225	.984	2.68	4	1	100	B11	received in	
-.401	.950	2.81	4	1	100	B12	the	
-.681	.803	3.11	4	1	100	B13	following	
-.536	.950	2.84	4	1	100	B14	domains	Classroom
-.233	.987	2.66	4	1	100	B15	by ticking	-focused
-.484	.899	2.86	4	1	100	B16	the box.	LTA
-.267	.869	2.85	4	1	100	B17		
-.077	1.059	2.48	4	1	100	B18		
-.069	.702	2.05	3	1	100	C1		
-.158	.686	2.12	3	1	100	C2	Please	
-.061	.640	2.07	3	1	100	C3	specify the	
-.027	.696	2.02	3	1	100	C4	amount of	
-.104	.668	2.09	3	1	100	C5	training	
-.186	.672	2.15	3	1	100	C6	you think	
-.132	.712	2.09	3	1	100	C7	you need	
-.098	.677	2.08	3	1	100	C8	in the	
-.091	.653	2.09	3	1	100	C9	following	
-.132	.689	2.10	3	1	100	C10	domains	

-.044	.717	2.03	3	1	100	C11	by ticking	
-.114	.706	2.08	3	1	100	C12	the box.	
-.249	.687	2.18	3	1	100	C13		
-.214	.714	2.14	3	1	100	C14		
-.158	.732	2.10	3	1	100	C15		
-.189	.706	2.13	3	1	100	C16		
-.075	.646	2.08	3	1	100	C17		
-.132	.712	2.09	3	1	100	C18		
-.349	.917	2.78	4	1	100	D1	Please	
-.349	.908	2.77	4	1	100	D2	specify the	
-.442	.938	2.78	4	1	100	D3	amount of	
-.371	.981	2.74	4	1	100	D4	training	
-.196	.977	2.66	4	1	100	D5	you think	
-.404	.924	2.71	4	1	100	D6	you have	
-.280	1.061	2.69	4	1	100	D7	received in	
-.291	.952	2.73	4	1	100	D8	the	
.032	.977	2.57	4	1	100	D9	following domains by ticking the box.	Knowledge of Testing and Assessment
-.057	.710	2.04	3	1	100	E1	Please	
-.075	.628	2.10	3	1	100	E2	specify the	
-.121	.674	2.10	3	1	100	E3	amount of	
-.066	.664	2.06	3	1	100	E4	training	
-.039	.650	2.04	3	1	100	E5	you think	
-.027	.696	2.02	3	1	100	E6	you need	
-.078	.730	2.05	3	1	100	E7	in the	
-.072	.655	2.07	3	1	100	E8	following	
-.072	.655	2.07	3	1	100	E9	domains by ticking the box.	
-.148	.943	2.60	4	1	100	F1	Please	
-.562	.869	2.95	4	1	100	F2	specify the	
-.248	.986	2.72	4	1	100	F3	amount of	
-.410	.910	2.80	4	1	100	F4	training	Purposes of Testing
-.568	.816	3.00	4	1	100	F5	you think	
-.461	.943	2.83	4	1	100	F6	you have	
-.311	.911	2.72	4	1	100	F7	received in	

							the following domains by ticking the box.	
-.022	.666	2.02	3	1	100	G1	Please specify the amount of training you think you need in the following domains by ticking the box.	
-.271	.692	2.19	3	1	100	G2		
-.104	.668	2.09	3	1	100	G3		
-.170	.700	2.12	3	1	100	G4		
-.177	.691	2.13	3	1	100	G5		
-.085	.609	2.15	3	1	100	G6		
-.098	.677	2.08	3	1	100	G7		
-.204	.959	2.70	4	1	100	H1	Please specify the amount of training you think you have received in the following domains by ticking the box.	Content and concepts of LTA
-.444	.957	2.71	4	1	100	H2		
-.349	.917	2.78	4	1	100	H3		
-.373	.927	2.70	4	1	100	H4		
-.457	.884	2.87	4	1	100	H5		
-.334	.880	2.85	4	1	100	H6		
-.117	.905	2.70	4	1	100	H7		
-.300	.880	2.75	4	1	100	H8		
-.354	.849	2.87	4	1	100	H9		
-.523	.853	2.86	4	1	100	H10		
-.621	.847	2.90	4	1	100	H11		
-.361	.863	2.89	4	1	100	H12		
-.515	.880	2.85	4	1	100	H13		
-.364	.976	2.76	4	1	100	H14		
-.245	.969	2.70	4	1	100	H15		
-.327	.902	2.79	4	1	100	H16		
-.240	.972	2.79	4	1	100	H17		
-.128	.918	2.79	4	1	100	H18		
-.248	.900	2.74	4	1	100	H19		
-.216	.937	2.59	4	1	100	H20		
-.079	1.009	2.55	4	1	100	H21		

-.239	.907	2.65	4	1	100	H22
-.075	.628	2.10	3	1	100	I1
-.052	.657	2.05	3	1	100	I2
-.177	.691	2.13	3	1	100	I3
-.047	.633	2.06	3	1	100	I4
-.163	.677	2.13	3	1	100	I5
-.151	.695	2.11	3	1	100	I6
-.197	.697	2.14	3	1	100	I7
-.339	.705	2.22	3	1	100	I8
-.235	.719	2.15	3	1	100	I9
-.163	.677	2.13	3	1	100	I10
-.281	.716	2.18	3	1	100	I11
-.132	.712	2.09	3	1	100	I12
-.206	.657	2.18	3	1	100	I13
-.179	.642	2.18	3	1	100	I14
-.206	.677	2.16	3	1	100	I15
-.187	.662	2.16	3	1	100	I16
-.126	.620	2.17	3	1	100	I17
-.061	.640	2.07	3	1	100	I18
-.251	.677	2.19	3	1	100	I19
-.072	.655	2.07	3	1	100	I20
-.164	.647	2.16	3	1	100	I21
-.126	.620	2.17	3	1	100	I22

Please specify the amount of training you think you need in the following domains by ticking the box.

As Table 4.3. shows, the average of most questions on the Language Assessment Literacy variable is high, and this is a reasonable estimate. Since the data set's standard deviation is close to zero, it also indicates that the data are close to the mean and have little distribution. According to the results of the tables, the skewness coefficients according to the table are in the range (2+ and -2), and the table related to the probability distribution curve shows this data in a normal way. According to the results of the tables, the skewness coefficients according to the table are in the range (2+ and -2), and the table related to the probability distribution curve shows this data as normal.

Criterion deviation: One scatter indicator that shows how far the average data is from the average. Given that the data set's standard deviation is close to zero, it indicates that the

data are close to the mean and have little scatter, While the standard deviation is large and smaller than the range (+2 and -2), which indicates a significant scatter of data.

Skewness: Skewness is equal to the third normalised torque. According to the results of the tables, the skewness coefficients according to the table are presented in the range (+2 and -2), and the table related to the probability distribution curve of these data shows normal.

4.2.2. Inferential analysis

Cronbach's Alpha test or reliability or reliability of research component localization tool is a statistical test used to test the reliability or reliability of research component localisation tool, which is designed in a spectrum, and the answers are multiple-choice. The table of reliable statistics shows the views and professional experience of language teachers on research variables and criteria using the Cronbach's Alpha test:

Table 4.4: Reliability Statistics of Research Variables and Criteria Using Cronbach's Alpha Test

Research criteria	Cronbach's Alpha	Cronbach's Alpha coefficient is the sum of the criteria
Reflective Teaching	0.879	
Classroom-focused LTA	0.932	
Knowledge of Testing and Assessment	0.912	0.972
Purposes of Testing	0.854	
Content and concepts of LTA	0.949	

The reliability statistics of the criteria related to the research variables using Cronbach's Alpha test indicate high reliability of the research components localization tool because the average of Cronbach's Alpha coefficients is 0.972, which is higher than 0.7. Moreover, it shows that teachers' opinions and professional experience are reliable and have good reliability. In fact, due to the parametric nature of statistical tests in the present study, the Pearson correlation coefficient is used to evaluate the effectiveness and variability between the indicators of research variables:

4.3. Investigating the Correlation Between Research Variables

In fact, in the present study, the Pearson correlation coefficient is used to investigate the effectiveness and influence of the research variablesIn Table 4.5, we examined the correlation between the research variables.

Table 4.5 Correlation between variables

Correlation between indicator		Reflective Teaching	Classroom-focused LTA	Knowledge of Testing and Assessment	Purposes of Testing	Content and concepts of LTA
Reflective Teaching	r	1	.469	.377	.467	.393
	p		.000	.000	.000	.000
Classroom-focused LTA	r	.469	1	.777	.744	.735
	p	.000		.000	.000	.000
Knowledge of Testing and Assessment	r	.377	.777	1	.703	.811
	p	.000	.000		.000	.000
Purposes of Testing	r	.467	.744	.703	1	.726
	p	.000	.000	.000		.000
Content and concepts of LTA	r	.393	.735	.811	.726	1
	p	.000	.000	.000	.000	

r= Pearson Correlation Coefficient

p=Level of Significance

The table above shows the correlation between the research variables. Since the correlation coefficient sign is the slope of the regression line, there is a positive and supportive relationship between the variables "Reflective Teaching" and "Classroom-focused LTA" as research indicators (sig value less than 0. 5) because the Pearson correlation coefficient between them is equal to 0.469. On the other side, there is a positive and supportive relationship between "Reflective Teaching" and "Knowledge of Testing and Assessment" (sig value is less than 0.5) because the Pearson correlation coefficient between

them is equal to 0.377. There is a positive and supportive relationship between "Reflective Teaching" and "Purposes of Testing" (sig value is less than 0.5) because the Pearson correlation coefficient between them is 0.467. There is a positive and supportive relationship between "Reflective Teaching" and "Content and concepts of LTA" (sig value is less than 0.5) because the Pearson correlation coefficient between them is 0.393.

There is a positive and supportive relationship between "Classroom-focused LTA" and "Knowledge of Testing and Assessment" (sig value is less than 0.5) because the Pearson correlation coefficient between them is 0.777. There is a positive and supportive relationship between "Classroom-focused LTA" and "Purposes of Testing" (sig value is less than 0.5) because the Pearson correlation coefficient between them is equal to 0.744. There is a positive and supportive relationship between the variables "Classroom-focused LTA" and "Content and concepts of LTA" as research indicators (sig value is less than 0.5) because the Pearson correlation coefficient between them is equal to 0.735, and in The sum of all variables is highly correlated.

There is a positive and supportive relationship between "Knowledge of Testing and Assessment" and "Purposes of Testing" (sig value is less than 0.5) because the Pearson correlation coefficient between them is equal to 0.703. There is a positive and supportive relationship between the variables "Knowledge of Testing and Assessment" and "Content and concepts of LTA" as research indicators (sig value is less than 0.5) because the Pearson correlation coefficient between them is equal to 0.811 and In general, all variables are highly correlated.

There is a positive and supportive relationship between the variables "Purposes of Testing" and "Content and concepts of LTA" as research indicators (sig value is less than 0.5) because the Pearson correlation coefficient between them is equal to 0.726. In total, All variables are highly correlated.

4.3.1.   Testing research questions (hypotheses)

4.3.1.1. Convergent validity

Convergent validity, which assesses the degree of correspondence between each structure and its questions, is the second criterion for evaluating the fit of measurement models (indicators).

Table 4.6 Convergent validity results of latent research variables

Concealed variables	Mean extraction variance (AVE> 0.5)
Reflective Teaching	0.662
Classroom-focused LTA	0.594
Knowledge of Testing and Assessment	0.631
Purposes of Testing	0.672
Content and concepts of LTA	0.722

This criterion has been accepted as an appropriate value for latent variables in light of the fact that it is equal to 0.5 for AVE and in accordance with the results of the above table, so proving the acceptability of the convergent validity of the study.

4.3.1.2. Divergent Validity (Fornell and Locker)

Table 4.7 Correlation matrix and divergent validity study

Concealed variables	Reflective Teaching	Classroom-focused LAT	Knowledge of Testing and Assessment	Purposes of Testing	Content and concepts of LTA	Root (AVE)
Reflective Teaching	1					0.822335
Classroom-focused LTA	0.535813	1				0.773885
Knowledge of Testing and Assessment	0.639901	0.773717	1			0.710151
Purposes of Testing	0.648055	0.848025	0.825636	1		0.836985
Content and concepts of LTA	0.857419	0.784122	0.854215	0.80142	1	0.847574

Divergent validity was also measured by comparing the AVE root with the correlation between latent variables, and for each of the reflective structures of the AVE, the root should be greater than the correlation of that structure with other structures in the model. Therefore,

the software outputs in the table indicate that the measurement tool has the appropriate validity.

4.3.2. Structural model fit

4.3.2.1. Significance coefficients Z (values of t-) and results of direct relations

According to the table, because the coefficients of t- are high for most of the research variables, at the 95% confidence level, the significance of most of them is confirmed, and in the Table 4.8, the values of significant coefficients and path coefficients are discussed.

Table 4.8 Direct relationship results and significance coefficients of model questions

Research Questions	T- values	Path coefficient	Significance (sig)	Test result
How do EFL teachers consider themselves proficient writing instructors?	3.840	0.724	0.000	Acceptable
What do EFL teachers think of scoring accuracy when it comes to writing assessment?	5.069	0.304	0.000	Acceptable
What do EFL teachers think of general assessment when it comes to writing classrooms?	6.946	0.421	0.000	Acceptable
What do EFL teachers think of different writing assessment methods?	3.206	0.317	0.000	Acceptable

According to the fit of the structural model of the research, several criteria have been used, the first and most basic of which are the coefficients of significance z and the values of t. Fitting the structural model using t-values showed that the values were greater than 1.96 and were significant at the 95% confidence level.

4.3.2.2. R Squares or R2 criteria

It is a measure that shows the effect of an exogenous variable on an endogenous variable, and three values of 0.19, 0.33, and 0.67 are considered the criterion values for weak, medium, and robust values of R2.

Table 4.9: Results of R2 criterion for endogenous structures

Concealed variables	R^2
Reflective Teaching	0.652
Classroom-focused LTA	0.745
Knowledge of Testing and Assessment	0.524
Purposes of Testing	0.452
Content and concepts of LTA	0.658

The value of R2 has been calculated for the endogenous structures of the research, which confirms the structural model's suitability according to the criterion's three values.

4.3.2.3. Criterion Q2

This criterion determines the model's predictive power, and if the value of Q2 for an endogenous structure achieves three values of 0.02, 0.5 and 0.35, it indicates the weak, medium and strong predictive power, respectively. It has a related structure or structures.

Table 4.10: Q2 benchmark results for endogenous structures

Concealed variables	Q2
Reflective Teaching	0.067671
Classroom-focused LTA	0.297025
Knowledge of Testing and Assessment	0.079846
Purposes of Testing	0.658542
Content and concepts of LTA	0.496855

The results of the above table show the model's appropriate predictive power regarding the research's endogenous structures and confirm the structural model's appropriate value.

4.3.2.4. Criterion F2

Another factor that is considered in evaluating the model's validity is Cohen's effect size or F2. Impact size determines whether an independent latent variable significantly affects a dependent variable.

Table 4.11: Results of F2 criterion for the relationship between endogenous and exogenous structures

Concealed variables	F^2
Reflective Teaching	0.516
Classroom-focused LTA	0.526
Knowledge of Testing and Assessment	0.618
Purposes of Testing	0.612
Content and concepts of LTA	0.316

This value is calculated from the coefficient of determination R2. F2 values between 0.02 and 0.15 indicate a weak effect, between 0.15 and 0.35 indicate a moderate effect, and greater than 0.35 indicates a significant effect of the independent variable on the dependent variable.

4.3.3. Fit the overall model

4.3.3.1. Communalities

The value of this criterion is obtained from the average of the expected values of the hidden variables of the research.

Table 4.12: Communality rate and R2 of research variables

Concealed variables	R2	Communality
Reflective Teaching	0.652	0.676229
Classroom-focused LTA	0.745	0.658456
Knowledge of Testing and Assessment	0.524	0.598902
Purposes of Testing	0.452	0.504315
Content and concepts of LTA	0.658	0.700544

Communality = This value is obtained from the mean squared of the factor loads of each variable.

4.3.3.2. GOF criteria

To evaluate the fit of the general model, the GOF criterion is used, which three values of 0.01, 0.25 and 0.36 are introduced as weak, medium and strong values for GOF.

Table 4.13 General Model Fit Results

GOF	R2	Communality
0.699	0.562	0.741

According to the values obtained for GOF in the amount of 0.699, the fit of the general model is confirmed.

4.4. Discussion

According to the results, it can be said that teachers have been more careful in developing evaluations. Previously, it was thought that the only application of classroom assessment was to identify and rank students' learning based on success. Today, assessment is used to gauge students' understanding of course objectives and to inform and provide guidance. Assessed literacy teachers can understand the adverse effects of poor assessment on students' learning and motivation. Therefore, to properly assess learners, teachers must have an appropriate level of assessment literacy. Use the information to establish immediate and long-term goals.

Assessment literacy has been scientifically studied in various educational contexts. For example, we review several studies on assessment literacy as follows.

In Webb's (2002) study, assessing student competence and performance was described as educators' assessment literacy. He also stated that extensive educational tricks could be used in the student classroom to encourage learning through self-assessment and peer assessment, as well as continuous descriptive feedback and assessment standards. In this regard, it showed that teachers who had passed the training course and assessment units had achieved significantly higher scores than other teachers.

Mertler (2005) showed that teachers are more evaluated than pre-service teachers while serving at the literacy level. Therefore, it can be said that an essential aspect of teacher

adequacy is strengthening assessment literacy or organizing assessments to translate learning objectives into assessment activities that confirm students' understanding and success. Thus, the lack or weakness of assessment knowledge raises concerns about how teachers are prepared to assess students. Teachers should organize classroom assessments in line with national standards because there is a relationship between the quality of classroom assessments and students' performance in standardized assessments.

Therefore, knowledge about different assessment types allows teachers to choose the most appropriate and practical tools to achieve learning objectives. Assessment literacy is, therefore, part of the professional identity of teachers, so those who have been educated and trained as language teachers are expected to accompany the currently accepted curriculum in the field of testing and assessment. So, what distinguishes general language teachers from English teachers for particular purposes is that they must understand their work's specific purposes, importance, and necessity and be willing to comply with those needs. Special-purpose English teachers are expected to help learners develop and improve the skills necessary to understand, apply, and present valid information in their work.

Lack of knowledge and expertise in teacher evaluation can damage students' time, motivation, and self-esteem and poor evaluation techniques. Because teachers devote a significant amount of their professional time to assessment-related tasks, they often cannot do it effectively. Language assessment literacy (LAL) is considered a fundamental construct in language assessment literature.

From the sociocultural view of learning, teachers' assessment literacy is a dynamic process that brings together teachers' assessment knowledge, assessment skills, and conceptions of assessment concerning their contexts of practice.

Since teacher literacy has been increasingly considered in recent years in programs related to teacher professional development, assessment and assessment literacy can help teachers and students understand the essential role of assessment in the classroom environment, applying the meaningful and continuous assessment and assessment not only for teachers and students but also for all education actors. It is necessary and valuable. Poor teacher education in the field of assessment and evaluation causes a qualitative decline in learning indicators in students.

5. CONCLUSION

5.1. Overview of the Study

The final chapter of the current study summarizes the contents of the previous chapters and concluding. Practical solutions for future research are presented. After presenting the research achievements and the obtained results, suggestions are presented based on the findings obtained during this research. These suggestions can be solutions to the community-level problem presented to those who are the target of this research.

5.2. Conclusion

According to the table, because the coefficients of t are high for most of the research variables, at the 95% confidence level, the significance of most of them is confirmed, and in the Table 5.1, the values of significant coefficients and path coefficients are discussed.

Table 5.1 Direct relationship results and significance coefficients of model questions

Research Questions	T- values	Path coefficient	Significance (sig)	Test result
How do EFL teachers consider themselves proficient writing instructors?	3.840	0.724	0.000	Acceptable
What do EFL teachers think of scoring accuracy when it comes to writing assessment?	5.069	0.304	0.000	Acceptable
What do EFL teachers think of general assessment when it comes to writing classrooms?	6.946	0.421	0.000	Acceptable
What do EFL teachers think of different writing assessment methods?	3.206	0.317	0.000	Acceptable

According to the fit of the structural model of the research, several criteria have been used, the first and most basic of which are the coefficients of significance z and the values of t. Fitting the structural model using t-values showed that the values were greater than 1.96 and were significant at the 95% confidence level.

According to R2, it has been calculated for endogenous structures of the research, which confirms the suitability of the structural model according to the three criteria. Criterion 2 Q indicates the appropriate predictive power of the model regarding the research's endogenous structures and confirms the structural model's appropriate value. Criterion Results showed an appropriate value for the relationship between endogenous and exogenous structures. GOF criterion is used to check the fit of the general model. According to the results, GOF = 0.699, which is a substantial value, and the fit of the general model is confirmed.

5.3. Pedagogical Implications

In order to achieve perfection, the reflective teaching style requires asking oneself the following questions: Which educational model am I employing? What relevance does it have to certain instructional contexts? How well does it function? To make the classroom into a form of laboratory where the teacher may connect teaching theory to actual teaching practice, teacher educators should put this theory into practice in the classroom. They should then examine the results and reflect on them. The teacher educator inspires teacher candidates by using reflective teaching methodology, allowing them to put it into practice during their teaching practice sessions. The thesis's main point is the value of reflective teaching methodologies to maintain interested teachers. In addition to the first topic which was about reflective teaching, language assessment literacy is also valuable for teachers as pedagogically. According to Scarino (2013), it is important to consider instructors' interpretive frameworks in addition to knowledge, abilities, and concepts in LAL. To put it another way, conversations in LAL must take into account the unique circumstances, methods, attitudes, and ideas that language teachers use in their classrooms. In order to promote professional development, it is especially crucial to acknowledge the interpretive frameworks of language teachers. Teachers' ways of thinking and behaving regarding the act of assessment coexist with their knowledge, abilities, and principles in language evaluation. As a result, Scarino clarifies how the elements of a language teacher's LAL interact with one another, a concept that has been briefly discussed by other authors (Fulcher, 2012; Taylor, 2009).

5.4. Suggestions for Future Research

- Researchers also suggest using other structural equation software in their research.
- It is suggested that, if possible, research should be done according to other community variables.
- It is suggested to combine these research variables with other influential variables in the field of literacy assessment and evaluation of teachers in education.
- It is suggested that for larger samples and in different communities with different strategies, this research be repeated by other researchers to increase the generalisability of the results.
- It is suggested that other researchers examine the performance of teacher literacy assessment in teacher education in larger communities and with respect to other influential variables.

5.5. Research Suggestions

Finally, the scientific uses of this research and suggestions for other researchers to continue the work will be expressed. One of the advantages of using a case study in research is that a case study is a systematic way of reviewing events, collecting data, analyzing information, and reporting results for "Exploring The Relationship Between EFL Teachers 'Language Assessment Literacy and Their Reflective Teaching", Using comments and people.

5.5.1. Practical suggestions

- It is suggested that education develop strategies to develop its teacher evaluation literacy.
- It is suggested that emphasis be placed on providing evaluation training to improve assessment literacy for all future English teachers.
- It is suggested that teachers be shown ways to become literate in the field of language assessment.
- It is suggested to provide methods through which test experts and teacher trainers can achieve this critical goal.

- It is suggested that teacher literacy assessment in education be increasingly considered in programs related to teacher professional development.

- It is suggested that by supporting measures that can improve and enhance teachers' knowledge of assessment and evaluation and a better understanding of assessment and evaluation, emphasis should be placed on facilitating the learning process.

- It is suggested that research on facilitating different levels in different areas related to the development of assessment literacy should be supported. Not only successful but also unsuccessful cases of in-service teachers should be considered as an effective and path-breaking way to increase teachers' awareness of their environment.

5.6. Main Research Findings

As can be seen from the opinions and experiences of the people, the average is almost well estimated. According to the results of the tables, the skewness coefficients according to the table are presented in the range (+2 and 2-), and the table related to the probability distribution curve of these data shows normal. The results of reliable statistics of criteria related to the sum of research variables using Cronbach's alpha test indicate high reliability. According to the table above, the average of Cronbach's alpha coefficients for the set of criteria is estimated to be 0.972, which shows a reasonable degree of reliability because it is calculated above 0.7. Correlation is observed between research variables. Since the sign of correlation coefficient is the slope of the regression line, there is a positive and significant relationship between the variables "Reflective Teaching" and "Classroom-focused LTA" as research indicators (sig value is less than 0.5)., Because the Pearson correlation coefficient between them is equal to 0.469.

On the other hand, there is a positive and significant relationship between "Reflective Teaching" and "Knowledge of Testing and Assessment" (sig value is less than 0.5) because the Pearson correlation coefficient between them is equal to 0.377. There is a positive and significant relationship between "Reflective Teaching" and "Purposes of Testing" (sig value is less than 0.5) because the Pearson correlation coefficient between them is 0.467. There is a positive and significant relationship between "Reflective Teaching" and "Content and concepts of LTA" (sig value is less than 0.5) because the Pearson correlation coefficient between them is 0.393.

There is a positive and significant relationship between "Classroom-focused LTA" and "Knowledge of Testing and Assessment" (sig value is less than 0.5) because the Pearson correlation coefficient between them is 0.777. There is a positive and significant relationship between "Classroom-focused LTA" and "Purposes of Testing" (sig value is less than 0.5) because the Pearson correlation coefficient between them is equal to 0.744. There is a positive and significant relationship between the variables "Classroom-focused LTA" and "Content and concepts of LTA" as research indicators (sig value is less than 0.5) because the Pearson correlation coefficient between them is equal to 0.735, and in The sum of all variables is highly correlated.

There is a positive and significant relationship between "Knowledge of Testing and Assessment" and "Purposes of Testing" (sig value is less than 0.5) because the Pearson correlation coefficient between them is equal to 0.703. There is a positive and significant relationship between the variables "Knowledge of Testing and Assessment" and "Content and concepts of LTA" as research indicators (sig value is less than 0.5) because the Pearson correlation coefficient between them is equal to 0.811 and In general, all variables are highly correlated.

There is a positive and significant relationship between the variables "Purposes of Testing" and "Content and concepts of LTA" as research indicators (sig value is less than 0.5) because the Pearson correlation coefficient between them is equal to 0.726. In total, all variables are highly correlated.

According to the table, the significance of most of them is confirmed because the coefficients of t for most of the research variables are high at the 95% confidence level. According to the values obtained for GOF in the amount of 0.699, the fit of the general model is confirmed.

REFERENCES

Alger, C. (2006). 'What went well, what didn't go so well': Growth of reflection in pre-service teachers. *Reflective practice*, *7*(3), 287-301.

Amobi, F. A. (2003). Finding and speaking their own voices: Using an online survey to elicit pre-service teachers' reflectivity about educational beliefs. *Reflective Practice*, *4*(3), 345-360.

Akbari, R., Behzadpoor, F., & Dadvand, B. (2010). *Development of English language teaching reflection inventory. System, 38*, 211–227.10.1016/j.system.2010.03.003

Bachor, D. G., & Baer, M. R. (2001). *An examination of preservice teachers' simulated classroom assessment practices*. Alberta journal of educational research, 47(3), 244.

Bailey, K. M. (2010). Observing classroom lessons for professional development. In: G. Park, H. P. Widod, and A. Cichocki (Ed) *Observing of teaching: Bridging theory and practice through research on teaching* (pp. 19-35). Munich, Germany: Lincom Europa.

Boyles-Baise, M. (2005). *Preparing community-oriented teachers: Reflections from a multicultural service-learning project.* Journal of Teacher Education, 56(5), 446-458.

Brookfield, S. D. (1995). *Becoming a critically reflective teacher.* San Francisco: Jossey-Bass.

Burhan-Horasanlı, E., & Ortaçtepe, D. (2016). Reflective practice-oriented online discussions: A study on EFL teachers' reflection-on, in and for-action. *Teaching and Teacher Education*, *59*, 372-382.

Campbell, C., & Evans, J. A. (2000). *Investigation of preservice teachers' classroom assessment practices during student teaching.* The Journal of Educational Research, 93(6), 350-355.

Campbell, C., Murphy, J. A., & Holt, J. K. (2002*). Psychometric analysis of an assessment literacy instrument: Applicability to preservice teachers.* In Annual meeting of the mid-western educational research association, Columbus, OH.

Cheng, L. (2008). The key to success: English language testing in China. *Language Testing*, *25*(1), 15-37.

Cizek, G. J. (2000). *Pockets of resistance in the assessment revolution.* Educational Measurement: Issues and Practice, 19(2), 16-23.

Clarke, A. (1995). Professional development in practicum settings: Reflective practice under scrutiny. *Teaching and teacher education*, *11*(3), 243-261.

Cole, A. L. (1997). *Impediments to reflective practice: Toward a new agenda for research on teaching.* Teachers and teaching, 3(1), 7-27.

Cornford, I. R. (2002). *Reflective teaching: Empirical research findings and some implications for teacher education.* Journal of Vocational education and Training, 54(2), 219-236.

Coyle, D. (2002). *The Case for Reflective Model of Teacher Education in Fundamental Principles.* Module Ed.D School of Education University of Nottingham.

Crusan, D. J. (2016). Writing Assessment Literacy: What Teachers Need to Know.

Crusan, D., Plakans, L., & Gebril, A. (2016). Writing assessment literacy: Surveying second language teachers' knowledge, beliefs, and practices. *Assessing writing*, *28*, 43-56.

Davison, C., & Leung, C. (2009). Current issues in English language teacher-based assessment. *TESOL quarterly*, *43*(3), 393-415.

Davidheiser, S. A. (2013). *Identifying areas for high school teacher development: A study of assessment literacy in the Central Bucks School District (Doctoral dissertation),* Drexel University).

DeLuca, C., & Klinger, D. A. (2010). *Assessment literacy development: Identifying gaps in teacher candidates' learning.* Assessment in Education: Principles, Policy & Practice, 17(4), 419-438.

DeLuca, C., & Johnson, S. (2017). Developing assessment capable teachers in this age of accountability. *Assessment in Education: Principles, Policy & Practice*, *24*(2), 121-126.

Dempsey, M. S., PytlikZillig, L. M., & Bruning, R. H. (2009). Helping preservice teachers learn to assess writing: Practice and feedback in a Web-based environment. *Assessing writing*, *14*(1), 38-61.

Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the education process.* Boston: DC. Heath and Company.

Donaghue, H. (2003). An instrument to elicit teachers' beliefs and assumptions. *ELT journal*, *57*(4), 344-351.

Earl, L. M., & Katz, S. (Eds.). (2006). *Leading schools in a data-rich world: Harnessing data for school improvement.* Corwin Press.

EFL teachers. Hacettepe University Journal of Education, 34(3), 602-620.

El-Dib, M. A. B. (2007). *Levels of reflection in action research. An overview and an assessment tool.* Teaching and teacher education, 23(1), 24-35.

Farrell, T. S. C. (2004). *Reflective practice in action*. Thousand Oaks, CA: Corwin Press.

Farrell, T. S. C. (2010). *Professional development through reflective practice in and for action*. In: G. Park, H. P. Widodo, and A. Cichocki (Ed), *Observation of teaching: Bridging theory and practise through research on teaching.* Munich, Germany: Lincom Europa, pp. 37-47.

Farrell, T. S. (2012). Reflecting on Reflective Practice:(Re) Visiting Dewey and Schon. *Tesol Journal*, *3*(1), 7-16.

Farrell, T. S. (2012). Novice-service language teacher development: Bridging the gap between preservice and in-service education and development. *Tesol Quarterly*, *46*(3), 435-449.

Farrell, T. S. C. (2015). *Promoting teacher reflection in second language education*. New York: Routledge.

Farrell, T. S. (2016). Anniversary article: The practices of encouraging TESOL teachers to engage in reflective practice: An appraisal of recent research contributions. *Language Teaching Research*, *20*(2), 223-247.

Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, *9*(2), 113-132.

Galluzzo, G. R. (2005). *Performance assessment and renewing teacher education the possibilities of the NBPTS standards.* The Clearing House: A Journal of Educational Strategies, Issues and Ideas, 78(4), 142-145.

Gobena, G. A. (2016). *Attitude of school principals, supervisors and mentees towards action research as reflective practices.* International Journal of Instruction, 10(1).

Griffiths, V. (2000). *The reflective dimension in teacher education.* International Journal of Educational Research, 33(5), 539-555.

Hamp-Lyons, L., & Heasley, B. (2006). *Study writing: A course in written English for academic purposes*. Cambridge University Press.

Harding, L., & Kremmel, B. (2016). 25. Teacher assessment literacy and professional development. In *Handbook of second language assessment* (pp. 413-428). De Gruyter Mouton.

Hawe, E. M., & Dixon, H. R. (2014). Building students' evaluative and productive expertise in the writing classroom. *Assessing Writing*, *19*, 66-79.

Hillier, Y. (2005). *Reflective teaching in further and adult education*. A&C Black.

Hyland, F. (1998). The impact of teacher written feedback on individual writers. *Journal of second language writing*, *7*(3), 255-286.

Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*, *25*(3), 385-402.

Jay, J. K., & Johnson, K. L. (2002). *Capturing complexity: a typology of reflective practice for teacher education.* Teaching and Teacher Education, 18, 73-85.

Jin, Y. (2010). The place of language testing and assessment in the professional preparation of foreign language teachers in China. *Language Testing*, *27*(4), 555-584.

Johnson, K., & Brumfit, C. (1979). *The communicative approach to language teaching*. Oxford University Press.

Larrivee, B. (2000). *Transforming teaching practice: Becoming the critically reflective teacher.* Reflective practice, 1(3), 293-307.

Larrivee, B. (2008). Development of a tool to assess teachers' level of reflective practice. *Reflective practice*, *9*(3), 341-360.

Leitch, R., & Day, C. (2000). *Action research and reflective practice: Towards a holistic view.* Educational Action Research, 8(1), 179-193.

Marcos, J. M., Sanchez, E., & Tillema, H. H. (2011). Promoting teacher reflection: What is said to be done. *Journal of Education for Teaching*, *37*(1), 21-36.

McMillan, J. H. (2000). *Fundamental assessment principles for teachers and school administrators.* Practical Assessment, Research & Evaluation, 7(8), 1-9.

Mede, E., & Atay, D. (2017). English language teachers' assessment literacy: The Turkish context.

Mertler, C. A. (2004). *Secondary teachers' assessment literacy: Does classroom experience make a difference.* American Secondary Education, 49-64.

Mertler, C. A., & Campbell, C. (2005). Measuring Teachers' Knowledge & Application of Classroom Assessment Concepts: Development of the" Assessment Literacy Inventory". *Online Submission*.

Mertler, C. A. (2009). Teachers' assessment knowledge and their perceptions of the impact of classroom assessment professional development. *Improving schools*, *12*(2), 101-113.

Odo, D. (2015). *Improving urban teachers' assessment literacy through synergistic individualized tutoring and self-reflection.* Networks: An Online Journal for Teacher Research, 17(2), 1-13.

Pill, J., & Harding, L. (2013). Defining the language assessment literacy gap: Evidence from a parliamentary inquiry. *Language Testing*, *30*(3), 381-402.

Pollard, A. (2002). *Reflective teaching: Effective and evidence-informed professional practice* ([New ed.], Andrew Pollard with contributions by Janet Collins. London, England: Continuum.

Pollard, A., & Collins, J. (2005). *Reflective teaching*. A&C Black.

Pollard, V. (2008). Ethics and reflective practice: Continuing the conversation. *Reflective Practice*, *9*(4), 399-407.

Popham, W. J. (2001). Teaching to the Test?. *Educational leadership*, *58*(6), 16-21.

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

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

Popham, W. J. (2010). *Everything school leaders need to know about assessment.* Thousand Oaks: Corwin Press.

Qing, X. U. (2009). Reflective Teaching--an Effective Path for EFL Teacher's Professional Development. *Canadian Social Science*, *5*(2), 35-40.

Qian, D. D. (2014). School-based English language assessment as a high-stakes examination component in Hong Kong: Insights of frontline assessors. *Assessment in Education: Principles, Policy & Practice*, *21*(3), 251-270.

Rein, M., & Schön, D. (1996). Frame-critical policy analysis and frame-reflective policy practice. *Knowledge and policy*, *9*(1), 85-104.

Richards, J. C., & Lockhart, C. (1996). *Reflective teaching in second language classrooms.* TESOL Journal, 1(2), 7-10.

Richards, Jack C., & Lockhart, Charles, *Reflective Teaching In Second Language Classrooms.* Cambridge: Cambridge University Press, 1994, Pp. 218 (Cambridge Language Education).

Rose, M. A. (2007). *Perceptions of technological literacy among science, technology, engineering, and mathematics leaders.*

Schön, D. A. (1987). *Educating the reflective practitioner: Toward a new design for teaching and learning in the professions.* Jossey-Bass.

Soltanpour, F., & Valizadeh, M. (2019). Iranian EFL Teachers' Writing Assessment Beliefs, Literacy, and Training Needs: Do Majors Matter? *Journal on English Language Teaching*, *9*(2), 26-41.

Stanley, C. (1998). A framework for teacher reflectivity. *TESOL quarterly*, *32*(3), 584-591.

Stiggins, R. J. (1999). *Assessment, student confidence, and school success.* The Phi Delta Kappan, 81(3), 191-198.

Stiggins, R. J. (2002). *Assessment crisis: The absence of assessment for learning.* Phi Delta Kappan, 83(10), 758-765.

Stoynoff, S., Chapelle, C. A., & Viii, P. (2005). *TESOL tests and testing.* Alexandria, VA: TESOL.

Taylor, L. (2009). Developing assessment literacy. *Annual Review of Applied Linguistics*, *29*, 21-36.

Vogt, K., & Tsagari, D. (2014). Assessment literacy of foreign language teachers: Findings of a European study. *Language Assessment Quarterly*, *11*(4), 374-402.

Volante, L., & Fazio, X. (2007). *Exploring teacher candidates' assessment literacy: implications for teacher education reform and professional development.* Canadian Journal of Education, 30(3), 749.

Wang, M. F., & Bakken, L. L. (2004). An academic writing needs assessment of English-as-a-second-language clinical investigators. *Journal of Continuing Education in the Health Professions*, *24*(3), 181-189.

Webb, N. L. (2002, April). Assessment literacy in a standards-based urban education setting. In *annual meeting of the American Educational Research Association, New Orleans*.

Zeichner, K. M., & Liston, D. P. (1996). *Reflective teaching: an introduction*. Mahwah, NJ: Lawrence Erlbaum Associates.

APPENDICES

APPENDIX-A: English Language Teaching Reflection Inventory

Items	Never	Rarely	Sometimes	Often	Always
1) I have a file where I keep my accounts of my teaching for reviewing purposes.					
2) I talk about my classroom experiences with my colleagues and seek their advice/feedback					
3) After each lesson, I write about the accomplishments/failures of that lesson or I talk about the lesson to a colleague					
4) I discuss practical/theoretical issues with my colleagues.					
5) I observe other teachers' classrooms to learn about their efficient practices					
6) I ask my peers to observe my teaching and comment on my teaching performance.					
7) I read books/articles related to effective teaching to improve my classroom performance.					
8) I participate in workshops/conferences related to teaching/learning issues.					
9) I think of writing articles based on my classroom experiences.					
10) I look at journal articles or search the internet to see what the recent developments in my profession are.					
11) I carry out small scale research activities in my classes to become better informed of learning/teaching processes.					
12) I think of classroom events as potential research topics and think of finding a method for investigating them.					

13) I talk to my students to learn about their learning styles and preferences.					
14) I talk to my students to learn about their family backgrounds, hobbies, interests, and abilities.					
15) I ask my students whether they like a teaching task or not.					
16) As a teacher, I think about my teaching philosophy and the way it is affecting my teaching.					
17) I think of the ways my biography or my background affects the way I define myself as a teacher.					
18) I think of the meaning or significance of my job as a teacher.					
19) I try to find out which aspects of my teaching provide me with a sense of satisfaction.					
20) I think about my strengths and weaknesses as a teacher.					
21) I think of the positive/negative role models that I have had as a student and the way they have affected me in my practice.					
22) I think of inconsistencies and contradictions that occur in my classroom practice.					
23) I think about the instances of social injustice in my own surroundings and try to discuss them in my classes.					
24) I think of ways to enable my students to change their social lives in fighting poverty, discrimination, and gender bias.					
25) In my teaching, I include less-discussed topics, such as old age, AIDS, discrimination against women andminorities, and poverty.					
26) I think about the political aspects of my teaching and the way I may affect my students' political views.					

27) I think of ways through which I can promote tolerance and democracy in my classes and in the society in general.					
28) I think about the ways gender, social class, and race influence my students' achievements.					
29) I think of outside social events that can influence my teaching inside the class.					

APPENDIX-B: Language Assessment Literacy Inventory

I. **Classroom-focused LTA**

    A. Please specify the amount of training you think you have received in the following domains by ticking the box.

	Training Received			
	None	Little (1-2 days)	Sufficient	Advanced
1)Preparing classroom tests				
2)Preparing diagnostic tests				
3) Preparing achievement tests				
4) Preparing proficiency tests				
5) Preparing placement tests				
6) Preparing progress				
7) Preparing language aptitude tests				
8)Using ready- made tests from textbook packages or from other				
9) Adapting ready-made tests for the				

10)Stages of language test construction (e.g. objectives, drawing up				
11)Scoring				
12)Grading				
13)Giving feedback to students based on information from				
14)Interpreting test				
15)Using self/peer				
16)Using informal, non-test type of assessment (e.g.essays, presentations,				
17)Using continuous type of assessment				
18)Using European Language Portfolio				

B. Please specify the amount of training you think you need in the following domains by ticking the box.

	Training		
	Needed		
	None	Yes, basic training	Yes, more advanced training
1)Preparing classroom tests			

2)Preparing diagnostic tests			
3) Preparing achievement tests			
4) Preparing proficiency tests			
5) Preparing placement tests			
6) Preparing progress tests			
7) Preparing language aptitude tests			
8)Using ready- made tests from textbook packages or from other sources			
9) Adapting ready-made tests for the needs of			
10)Stages of language test construction (e.g. objectives, drawing up test specifications...)			
11)Scoring			
12)Grading			
13)Giving feedback to students based on information from tests/assessment			
14)Interpreting test scores			
15)Using self/peer assessment			
16)Using informal, non- test type of assessment (e.g.essays, presentations, homeworks)			

17)Using continuous type of assessment (e.g.quizzes)			
18)Using European Language Portfolio			

II. Knowledge of Testing and Assessment

A. Please specify the amount of training you think you have received in the following domains by ticking the box.

	Training Received			
	None	Little (1-2 days)	Sufficient	Advanced
1)Informal/ Formal assessment				
2)Formative/Summative				
3)Norm/Criterion- referenced				
4)Discrete point/Integrative				
5)Direct/Indirect testing				
6)Objective/Subjective testing				
7)Approaches to language testing (e.g. integrative, communicative,				
8)Alternative assessment				
9)Computer-based testing				

B. Please specify the amount of training you think you need in the following domains by ticking the box.

	Training Needed		
	None	Yes, basic training	Yes, more advanced training
1)Informal/ Formal assessment			
2)Formative/Summative			
3)Norm/Criterion- referenced			
4)Discrete point/Integrative			
5)Direct/Indirect testing			
6)Objective/Subjective testing			
7)Approaches to language testing (e.g. integrative, communicative,			
8)Alternative assessment			
9)Computer-based testing			

III. Purposes of Testing

    A. Please specify the amount of training you think you have received in the following domains by ticking the box.

	Training Received			
	None	Little (1-2 days)	Sufficient	Advanced
1)Giving grades				
2)Finding out what needs to be learned/taught				
3)Placing students onto programs, courses, etc.				
4)Testing competence in a language				

5)Identifying what has			
6)Measuring general ability to learn a foreign			
7)Awarding final certificates (from school, program; local, regional/			

B. Please specify the amount of training you think you need in the following domains by ticking the box.

	Training		
	Received		
	None	Yes, basic training	Yes, more advanced training
1)Giving grades			
2)Finding out what needs to be learned/taught			
3)Placing students onto programs, courses, etc.			
4)Testing competence in a language			
5)Identifying what has			
6)Measuring general ability to learn a foreign			
7)Awarding final certificates (from school, program; local, regional/			

IV. Content and concepts of LTA

A. Please specify the amount of training you think you have received in the following domains by ticking the box.

	Training Received			
	None	Little (1-2 days)	Sufficient	Advanced
1)Testing reading in English				
2)Different test items/task types to test reading in English				
3)Testing listening in English				
4)Different test items/task types to test listening in English				
5)Testing speaking in English				
6)Different test items/task types to test speaking in English				
7)Testing writing in English				
8)Different test items/task types to test writing in English				
9)Testing Grammar in English				
10)Different test items/task types to test grammar in English				
11)Testing Vocabulary in English				
12)Different test items/task types to test vocabulary in English				
13)Testing integrated language skills				
14)Testing pronunciation in English				
15)Different test items/question types to test pronunciation in English				
16)Practicality				

17)Reliability (e.g. rater, test administration, test, student-				
18)Validity (face, construct, criterion, content)				
19)Authenticity				
20)Washback				
21)Using statistics to study the quality of tests / assessment				
22)Alternatives in assessment (portfolios, conferences, interviews, observations, self/peer				

B. Please specify the amount of training you think you need in the following domains by ticking the box.

	Training Received		
	None	Little (1-2 days)	Sufficient
1)Testing reading in English			
2)Different test items/task types to test reading in English			
3)Testing listening in English			
4)Different test items/task types to test listening in English			
5)Testing speaking in English			
6)Different test items/task types to test speaking in English			
7)Testing writing in English			

8)Different test items/task types to test writing in English			
9)Testing Grammar in English			
10)Different test items/task types to test grammar in English			
11)Testing Vocabulary in English			
12)Different test items/task types to test vocabulary in English			
13)Testing integrated language skills			
14)Testing pronunciation in English			
15)Different test items/question types to test pronunciation in English			
16)Practicality			
17)Reliability (e.g. rater, test administration, test, student-related)			
18)Validity (face, construct, criterion, content)			
19)Authenticity			
20)Washback			
21)Using statistics to study the quality of tests/assessment			
22)Alternatives in assessment (portfolios, conferences, interviews, observations, self/peer assessment)			