



An Overview of Language Assessment Literacy Research within English Language Education Context

İngiliz Dili Eğitimi Bağlamında Yabancı Dilde Ölçme-Değerlendirme Okuryazarlığı Araştırmalarına Genel Bir Bakış

Aylin SEVİMEL-SAHİN* 

Gonca SUBASI** 

Received: 24 December 2018

Review Article

Accepted: 11 September 2019

ABSTRACT: Language assessment literacy (LAL) is the indication of language teachers' familiarity with assessment processes and ability to perform successful assessment procedures in foreign language education. In recent years, its importance has increased due to the developments in foreign language teaching. The research into this concept has already begun to define and conceptualize it. However, empirical research on LAL, especially about educational contexts, has not been much widespread. So, this systematic review mainly aims to present an overview of LAL concept within the scope of foreign language education. In particular, this review firstly illustrates its conceptual side by exemplifying the definitions, components and characteristics. Then, focusing on English language teachers, it evaluates the most up-to-date empirical studies conducted in EFL in-service or pre-service teaching contexts to show its place in foreign language assessment and the recent trends in LAL research. Lastly, it discusses the findings to suggest implications for practice and further study about LAL.

Keywords: language assessment literacy, foreign language education, English language teaching, foreign language assessment.

ÖZ: Yabancı dilde ölçme-değerlendirme okuryazarlığı, dil öğretmenlerinin ölçme-değerlendirme süreçlerine aşina olmasının ve yabancı dil eğitiminde başarılı ölçme-değerlendirme yöntemlerini uygulama becerisinin bir göstergesidir. Son yıllarda yabancı dil öğretimindeki gelişmeler nedeniyle bu kavramın önemi artmıştır. Bu kavrama yönelik yapılan araştırmalar, onu tanımlamaya ve kavramsallaştırmaya çoktan başlamış bulunmaktadır. Bununla birlikte, özellikle eğitim bağlamında yabancı dilde ölçme-değerlendirme okuryazarlığı üzerine yapılan ampirik çalışmalar çok yaygın değildir. Dolayısıyla, bu sistematik derlemenin temel amacı, yabancı dil eğitimi kapsamında yabancı dilde ölçme-değerlendirme okuryazarlığı kavramının genel bir değerlendirmesini sunmaktır. Özellikle, bu derleme öncelikle bu kavramın tanımları, bileşenleri ve karakteristik özelliklerini örneklerdirerek onun kavramsal yanlarını göstermektedir. Sonrasında bu derleme, İngilizce öğretmenlerine odaklanarak kavramın yabancı dilde ölçme-değerlendirmedeki yerini ve bu kavrama bağlı güncel araştırma eğilimlerini ortaya koymak için hizmet içi ve hizmet öncesi öğretmenlik bağlamlarında gerçekleştirilen en güncel ampirik araştırmaları değerlendirmektedir. Son olarak, bu derleme, söz konusu kavramın uygulama alanları ve onun hakkında daha çok araştırma önerilerinde bulunmak için bulguları tartışmaktadır.

Anahtar kelimeler: yabancı dilde ölçme-değerlendirme okuryazarlığı, yabancı dil eğitimi, İngiliz dili eğitimi, yabancı dil değerlendirmesi.

* Corresponding Author: Res. Asst. Dr., Anadolu University, Eskisehir, Turkey, aylinsevimmel@anadolu.edu.tr, <https://orcid.org/0000-0003-2279-510X>

** Asst. Prof. Dr., Anadolu University, Eskisehir, Turkey, goncas@anadolu.edu.tr, <https://orcid.org/0000-0001-7049-5940>

Citation Information:

Sevimel-Sahin, A., & Subasi, G. (2019). An overview of language assessment literacy research within English language education context. *Kuramsal Eğitim Bilim Dergisi [Journal of Theoretical Educational Science]*, 12(4), 1340-1364.

Introduction

Teaching and assessment are considered to be two interrelated components in an educational context. In other words, assessment is acknowledged to be an essential part of instruction because it gives feedback about the quality of teaching. Therefore, assessment is helpful to reveal whether teaching is useful or not, whether students learn the targeted behaviors/knowledge/skills, and the extent the objectives of a course is achieved, and so on (Herrera & Macias, 2015; Rogier, 2014). Correspondingly, teachers make decisions about content, materials, alternatives, and the like in order for better and effective teaching practices and instruction (Rea-Dickins, 2004). So, since assessment gives information about teaching, it simply guides, regulates and reinforces teaching. As a result, students are motivated to learn, and both teachers' and students' performances are improved. Therefore, it can be asserted that teachers have dual roles: being an instructor, and at the same time, an assessor (Inbar-Lourie, 2013a; Wach, 2012). However, to be a proper assessor is a difficult characteristic because it entails both teachers' applying assessment practices according to their teaching environment and teaching knowledge, and developing the understanding of the nature of assessment (Scarino, 2013). Moreover, when the component of language itself is added to this difficulty, it is even more complex for teachers to be a good assessor. Therefore, it can be argued that language teachers have three main responsibilities: being proficient in a foreign language, having the necessary knowledge and skills in language teaching pedagogy, and being able to assess students' language progress in order to make decisions about the teaching/learning process. Out of these, language assessment becomes prominent because it guides teachers how to make decisions about the learning/teaching process (Oz & Atay, 2017). To carry out an effective assessment process, language teachers must possess a certain qualification as a part of their professional competence; namely, language assessment literacy (LAL).

LAL basically refers to the familiarity with the processes or procedures of language assessment. That is, teachers' ability to utilize effectively their knowledge of assessment in their teaching practices. In this sense, they are expected to have the knowledge and skills of assessment procedures such as how to design a testing task, what the suitable language measurement tools are, how to administer them, how to interpret the results to improve teaching/learning, and how to identify the needs of language learners. They are also expected to know what the approaches and recent trends in assessment are, how to relate that knowledge into their own practice, what the main principles of testing and assessment are, and how to evaluate a course entirely regarding students, teachers themselves, teachers' own teaching, inputs and outputs of the instruction. In order to have such features, language teachers must be assessment literate.

Recently, LAL has gained much importance because there has been a shift from traditional testing notion (summative evaluation) to current assessment concept (formative assessment) as well as a change in the approaches of language teaching methodology and the expectations about the education worldwide (Tsagari & Vogt, 2017). Therefore, as emphasized, language teachers must employ various assessment procedures that are suitable for their foreign language teaching context in order to keep up with modern world's needs of foreign language education. Moreover, LAL is seen as the link between the quality of assessment and the achievement of students (Bayat &

Rezaei, 2015). Though there have been notable studies into LAL lately, the research into this field can be said to be too new and scarce (Fulcher, 2012). So, researchers need to study on this concept from different perspectives in order to enlighten and provide insights into this issue more.

In response to this need, the purpose of the current systematic review is to present an overview of what is known, what has been studied, what the missing parts are, and what the implications for language education might be about LAL within English language teaching/learning context.

Literature Review

Language Assessment Literacy (LAL): The Conceptual Framework

There has been an increasing attention to demonstrate the role of assessment in language teaching by means of investigating language assessment theory and practice. As Inbar-Lourie (2017) has argued, there have been changes in the approach of language assessment towards a more formative understanding accompanied by the developments occurred in the pedagogy of language teaching. In the same vein, there has been an expansion of standardized language testing around the world and some countries use language tests as their national policy (Fulcher, 2012). Furthermore, as in the change of language assessment notion towards formative assessment, there has been a shift to more sociocultural theories of learning framed within zone of proximal development and dynamic assessment in language teaching methodologies (Fulcher, 2012; Inbar-Lourie, 2008; Scarino, 2013). Thus, the consequences of all such changes have led to modifications in language assessment conceptions. It can be inferred that the responsibilities of language teachers have increased due to the fact that the importance of language assessment has been recognized.

The importance of LAL concept has been emphasized in research studies conducted in assessment field because this concept is considered to meet the needs of being competent language educators in the field of testing, assessment and evaluation as well as being an effective instructor in language teaching process (Buyukkarci, 2016). In this respect, theories and practices of language assessment have obtained a crucial role in language teaching/learning while helping teachers to develop their own LAL competencies. Therefore, most researchers have studied LAL to define, to construct its frameworks, to determine the characteristics of language teachers that have LAL, and to underline its significance within foreign language education. In this respect, there have been a number of review or conceptual articles related to LAL over the last years compared to empirical research studies, especially within foreign language learning/teaching contexts.

As far as this issue is taken into account, the current review firstly presents what is done in LAL research by discussing its meaning and background, and then, evaluates the empirical research studies carried out with English language teachers as to display its place in the field of language assessment.

Definitions of LAL. The basis of LAL dates back to 1990s when Stiggins (1995) first coined the term assessment literacy (AL) in general education. He defined AL as teachers' own understanding of assessment knowledge and procedures and the

influence of such knowledge in their practices. After that, the content of AL in education was formed in line with the increasing role of formative assessment in education. However, the content was not sufficient to appeal to specific teaching subject areas such as mathematics, science and language arts. Thereupon, researchers began to investigate AL from the perspective of different subject areas. Since foreign language teachers have a specific competency, which is language proficiency, content pedagogical knowledge of language teaching differs from other subjects. This difference makes language assessment distinctive from teacher general assessment literacy because both the teaching context and target learner group are different. So, language assessment requires peculiar features in its own sense including how to assess language skills such as speaking, writing, listening and reading, and language areas such as grammar, vocabulary and pronunciation within the scope of culturally situated language teaching environments. Therefore, the necessity of being language assessment literate to fulfill the needs of language teaching has been highly emphasized in the literature.

Basically, LAL refers to the familiarity with assessment procedures of the stakeholders concerned in this process such as teachers (testers), students (testees), administrators and other staff in education. However, as in language teaching/learning process, the most central stakeholder is language teachers (Giraldo, 2018; Rea-Dickins, 2004). Thus, teachers have the responsibility of assessing their language learners most of the time except for large-scale standardized tests that require experts of test-designers. Based on this view, most researchers have defined LAL from the perspective of language teachers. Therefore, sometimes LAL is referred as language teacher assessment literacy (LTAL).

Davies (2008) is among the researchers who first tried to define the concept of LAL and determine its components. According to his definition, LAL requires three types of features: the knowledge of language, context, and measurement; training in assessment skills and methodology like item writing and statistics; and the principles of language assessment like ethics and impact. Likewise, Inbar-Lourie (2008) has presented that LAL is “to have capacity to ask and answer critical questions about the purpose of assessment, about the fitness of the tool being used, about testing conditions, and about what is going to happen on the basis of the results” (p. 389). She has also indicated that LAL includes “the understanding the ‘what’ and performing the ‘how’ necessitates appreciation of the background and reasoning behind the actions taken, that is, ‘why’” (Inbar-Lourie, 2008, p. 390). So, the questions of what, how and why lead to the construction of LAL meaning. Besides, Taylor (2009) has stated that LAL means understanding the language assessment principles, being able to put these principles into practice, selecting or developing appropriate assessment tasks, collecting assessment data to interpret and evaluate the process, and making decisions about pedagogical process. Similarly, LAL refers to the knowledge, principles and skills of language assessment procedures inclusive of developing language test items, administering them, interpreting test scores, evaluating the whole language learning/teaching process, and the functions of assessment in a broader sense (Lam, 2014; Pill & Harding, 2013). By drawing attention to the relationship between theory and practice of language assessment, Vogt and Tsagari (2014) have defined LAL as “the ability to design, develop, and critically evaluate tests and other assessment procedures, as well as the

ability to monitor, evaluate, grade and score assessments on the basis of theoretical knowledge” (p. 377). On other hand, Fulcher (2012) studied on the learning needs of in-service language teachers in assessment by developing a survey instrument. He focused on knowledge, skills, principles and concepts of language assessment by adding another dimension called contexts that involve historical, social, political and philosophical frameworks to make sense of language assessment itself and its effect on society. Considering these, he provides an expanded definition of LAL as:

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

This is now accepted as the most comprehensive definition of LAL in the literature. However, Giraldo Aristizabal (2018) has underscored the view that LAL must comprise not only large-scale (standardized) but also small-scale (classroom-based) language assessment knowledge and skills in addition to the appropriate use of assessment ethically and in a fair way.

To sum up, the definitions of LAL given to date have more or less highlighted the main elements such as teachers’ knowledge, skills, principles, contexts, procedures and understanding of language assessment within the scope of language teaching education.

Components or dimensions of LAL. Researchers who are interested in the field of LAL have illustrated its components or dimensions in relation to their definitions. That is, they have tried to demonstrate what this concept constitutes by developing frameworks or conceptualizations. For example, Davies (2008) has discussed that LAL has three components: knowledge, skills, and principles. Knowledge of language assessment entails the knowledge of language description, context and measurement. Skills encompass practices of assessment such as designing language tests and interpreting results. Principles involve using appropriate tests, fairness and washback effect of assessment. Inbar-Lourie (2008; 2013b) has also approached LAL concept by three questions in order to classify the dimensions of LAL: what, how, why. ‘What’ reflects the content or behavior or learning outcome about language learning to be assessed; ‘how’ refers to the assessment procedures, tools, analysis and the like; ‘why’ includes the rationale or framework behind language assessment. Out of these dimensions, ‘what’ is prominent because it directly reflects the context of language teaching. Besides, Scarino (2013) has divided LAL into two categories: knowledge base and process base. Knowledge base means identifying disciplinary domains within language context boundaries so as to show how to develop and use the knowledge in practice. Process base refers to the developing phase that teachers construct their own concepts through their critical approaches towards language assessment. As for Fulcher (2012), as in his definition, he has concluded that LAL has three layers: practices at the bottom layer, principles in the middle layer and contexts at the top layer. Practices consist of knowledge, skills and abilities related to language assessment. Principles are the processes, principles and concepts of language

assessment. Contexts are made up of historical, social, political and philosophical frameworks that question the origins, reasons and impacts of language assessment. Fulcher (2012) has also underlined the fact that this type of dimensional LAL may not fit to other stakeholders except for teachers. In order to summarize all the models of LAL, Stabler-Havener (2018) presents an outline of LAL components in which she has proposed four models. The first one is 'A Five-Component, Professional Development Program Model' which uses certain standards to identify AL components. Inspired by this model, she has argued that Inbar-Lourie's (2008) 'what, how, why' dimensions reflect those standards. The second one is 'A Skills, Knowledge and Principles Model' that is attributed to Davies' (2008) LAL components. The third one is 'A Practices, Principles and Contexts Model' which refers to Fulcher's (2012) expanded definition of LAL. The last one is 'A LAL Stakeholder Profile Model' that takes into account other stakeholders in addition to teachers as supported by Pill and Harding (2013), and Taylor (2013). For instance, Taylor (2013) has divided stakeholder constituents into three layers: core, intermediary and peripheral. In the peripheral layer, there are policy makers and general public. In the intermediary layer, language teachers and course instructors are found. In the core layer, test makers and researchers take place. According to their own layer type, the extent of dimension about knowledge and other related topics about language assessment change. For classroom teachers, language pedagogy is more important than other constituents in language assessment procedure (see Taylor, 2013, Figure 2, p. 410 for more information).

To conclude, LAL shares relatively the same components or dimensions with different categorical names such as knowledge, process, skills, principles, abilities and contexts.

Characteristics or competencies of LAL. Researchers or area experts in language assessment/education have drawn conclusions to characterize what the features of LAL are and what kind of competencies reflect teachers' LAL development. In this sense, the first steps were taken by the American Federation of Teachers, National Council on Measurement in Education and National Education Association in 1990 by suggesting seven principles for teacher AL under the name 'The Standards for Teacher Competence in the Educational Assessment of Students' (Mertler, 2003). In fact, The Standards were generated before the term AL that did not exist until Stiggins mentioned in 1995. But The Standards have been accepted to reflect what the competencies of teachers with regard to assessment are in terms of general education, which have been used to describe teacher assessment literacy competencies. The principles proposed cover the following characteristics:

1. choosing appropriate assessment methods;
2. developing appropriate assessment methods;
3. administering, scoring, and interpreting the results of assessments;
4. using assessment results to make decisions;
5. developing valid grading procedures;
6. communicating assessment results;
7. recognizing unethical or illegal practices (Mertler, 2003, pp. 8-9).

Stiggins (1995) has also pointed out that assessment literates know and realize what sound and unsound assessment is. That is, all assessment literate stakeholders know what they assess, how to assess, why they assess and how to use assessment results effectively. Similarly, some researchers believe that not only classroom teachers but also other stakeholders in assessment process such as policy makers, test developers and learners themselves are important factors (Djoub, 2017; Pill & Harding, 2013;

Taylor, 2013). Hence, all such stakeholders need to be assessment literate in order for a better educational process. Nonetheless, teachers have still held the central role among other stakeholders in terms of language assessment (Giraldo, 2018).

Most researchers have identified the competencies of assessment literate teachers nearly the same regardless of any specific teaching subject area (Berry & O'Sullivan, 2016; Gotch & French, 2014; Huang & He, 2016; Khadijeh & Amir, 2015; Rogier, 2014). They have emphasized common characteristics of assessment literate teachers as understanding what a good assessment is, knowledge of how to assess with which tools according to appropriate learning goals, interpreting assessment results in order to make decisions about teaching as well as to enhance learning, being able to analyze one's own assessment practices, integrating suitable assessment procedures into learning/teaching context effectively, being familiar with alternatives in assessment, and applying assessment procedures ethically. Nevertheless, such characteristics reflect only general competencies of assessment literate teachers. To highlight the factor of language, Inbar-Lourie (2013b) has postulated that language assessment literate teachers must have additional characteristics such as proficiency in language classroom assessment, mastering language acquisition and learning theories and modifying them accordingly to their assessment practices, considering authentic language use and being aware of assessment dilemmas. She has also indicated that all these characteristics are individualized within teachers' own language teaching context. In another research, Inbar-Lourie (2017) has exemplified the importance of language in LAL by means of the Common European Framework of Reference (CEFR). CEFR integrates teaching and assessment in language education, and thus, has recommended alternative forms of language assessment such as portfolio and self-assessment. Therefore, in a way, CEFR can be assumed to contribute to LAL.

Likewise, Giraldo (2018) has supported the claim that language component should be underlined while discussing the competencies of language assessment literate teachers because language itself - its knowledge, use, and pedagogy- differentiates LAL from general AL characteristics. He has constructed a comprehensive list of LAL characteristics by expanding Davies' (2008) components of LAL, and he has suggested a total of sixty-six features considering three-dimensional structure of LAL as knowledge, skills and principles (see Giraldo, 2018, Table.1, p. 188 for more information). He has tried to emphasize the element of language nearly in all sub-competencies of language assessment literate teachers in his list. Therefore, it can be inferred that Giraldo (2018) has presented a comprehensive and detailed list of LAL characteristics.

As a conclusion, it is apparent that characteristics of assessment literate teachers are relatively common. But it should be noted that other categories highlighting the component of language have also been added in terms of language education thanks to the research conducted within the scope of LAL.

Purpose of the Review

LAL is considered to be an important element in foreign language education, and the research into this field has just commenced as Fulcher (2012) has asserted. There have been recent attempts to define this concept, to develop related framework, to identify the characteristics or to review the existing literature through mostly a

theoretical or conceptual point of view (e.g. Inbar-Lourie, 2013a; Popham, 2009; Rogier, 2014; Taylor, 2009). Yet, there are few empirical research studies investigating this concept from the perspective of teaching context (in-service vs pre-service) in the literature. Therefore, the major concern of this systematic review is to provide an overview of the research studies conducted with in-service and pre-service English language teachers in the context of foreign countries where English language is taught as a foreign language. To explore recent trends in LAL within the scope of context can shed light on what is known, what is done, what needs to be done to improve, what the gaps or challenges are, how the concept is practiced in classrooms and what can be recommended as implications in education regarding the concept of LAL.

All in all, the current systematic review firstly aims to introduce the concept of LAL by discussing its definitions, characteristics, components and importance in foreign language education by means of the studies conducted in the section ‘LAL: The Conceptual Framework’. Then, specifically, this article focuses on the research studies carried out in EFL contexts both with in-service and pre-service teachers between 2014-January and 2018-June (the last five years) in order to present a review of current state of LAL in such contexts at the teacher level by showing not only conceptual but also practical side of LAL. Lastly, this paper has the purpose of summarizing and synthesizing the most up-to-date research on LAL within foreign language education so as to provide insights into future directions and implications for further study.

Method

The present study employs systematic review as the research methodology. It is because it aims to explore the concept of LAL within English language teaching contexts including both in-service and pre-service educational environments in order to determine the current place of LAL within these contexts, and also to make implications for future studies.

Systematic review basically refers to review relevant literature in a particular field of study in a systematic process as comprehensive as possible to make meaning of information and contribution to the addressed questions or problems (Petticrew & Roberts, 2006). It is based on a specific problem or question to review the related literature systematically rather than just summarizing it (Petticrew & Roberts, 2006). It should be noted that although systematic review is used interchangeably with the terms research synthesis or research review and sometimes with meta-analysis or meta-synthesis, there is not a general agreement how they are different (Copper, Hedges, & Valentine, 2010).

Systematic review is mostly utilized “when a general overall picture of the evidence in a topic area is needed to direct future research efforts” (Petticrew & Roberts, 2006, p. 21) including quantitative, qualitative or any other eligible research studies as in the present study; and thus, it does not have to be only quantitative (as in meta-analysis) or qualitative (as in meta-synthesis) research. Its systematicity comes from the procedure used as a step-by-step process to review the literature and accordingly, to provide an overview by means of synthesizing the studies to enlighten the addressed topic. According to Petticrew and Roberts (2006), systematic review includes the steps of formulating question, problem, hypothesis or topic needed to be searched, determining the eligibility criteria as inclusion and exclusion criteria,

literature search through databases, and synthesizing evidence by means of narrative synthesis. In narrative synthesis, the organization and description of studies, the analysis of the findings, and the synthesis of the findings are carried out (Petticrew & Roberts, 2006). Therefore, certain steps were performed while organizing the content of this systematic review.

The Procedure of the Present Systematic Review

This systematic review utilizes some steps to review the relevant literature in line with the study purpose considering Petticrew and Roberts' (2006) proposed process.

First of all, the relevant research about language assessment literacy was investigated. To do so, certain well-known electronic databases specifically about education (ERIC, EBSCO and Google Scholar) were utilized. It should be noted that ERIC was a part of EBSCO while searching. Additionally, a national database called DergiPark/ULAKBİM, a Turkish database, was included because some articles could not be found via mentioned databases though Turkey has an EFL context.

Secondly, the keyword "language assessment literacy" was used with and without quotation marks to do the preliminary search. That keyword was applied to all sections such as abstract, keywords, content and title of the studies.

In the third step, some criteria (inclusion and exclusion criteria) were employed according to the purpose of this study to determine and choose relevant research studies. Those criteria could be specified as: Only accessible studies that were written in English language were considered. The category of education was selected. The most recent research about LAL was determined by putting check between the years 2014-January and 2018-June so that last 5 years of research would appear. Only peer-reviewed empirical articles were taken into consideration because one of the aims of this review was to show the place of LAL within practice. Therefore, books, book chapters, book reviews, reviews of literature, theoretical/review articles, conference papers and proceedings, theses/dissertations were excluded. Besides, topic relevancy (language assessment literacy) was concerned; that is, the studies with reading and writing skills used to describe merely literacy without assessment focus were not thought for this review. The articles of only teacher focused LAL were included since EFL teachers are the main stakeholders of language assessment, and teachers' background contexts were taken into account as pre-service and in-service teaching contexts. It should be underlined that only English subject matter as a foreign language was considered. All types of research designs such as qualitative, quantitative and mixed-methods designs were welcomed. Overlapping articles within the databases were also removed.

As a fourth step, the organization and exclusion of articles were carried out by two authors of this study. Accordingly, the articles which suited to the criteria (only English language based, teacher-focused, empirical, not review, consisting of pre-service and in-service teaching contexts, the most up-to-date) were selected and examined according to the determined eligibility criteria. In the first search, there were 758 articles regardless of the criteria (EBSCO: 135, ERIC: 34, Dergipark: 351; Google Scholar: 238); there were overlapping ones and some articles without the scope in this number, as well. After the elimination, the rest was 21 empirical research articles all of which were suitable to the selection criteria.

Finally, these suitable articles about EFL teachers' LAL within pre-service and in-service teaching contexts were evaluated in order to show the current state of LAL in ELT as well as to provide future directions about the topic by means of narrative synthesis under the themes in-service and pre-service EFL teaching contexts.

The Profile of the Research Articles Included

The research articles reviewed in this paper are made up of 21 empirical studies conducted in the last 5 years. The distribution of the articles across the publication years can be seen in Table 1. In the last two years, there has been an expansion of empirical articles within LAL research. Thereupon, it can be inferred that rather than writing only review or conceptual articles, researchers have begun to investigate LAL in practice. On the other hand, the studies conducted in in-service EFL context ($f=17$) outnumbered the studies in pre-service context ($f=4$). Therefore, there is a research gap in terms of the studies in pre-service EFL contexts where the development of LAL begins, so it is a critical stage.

Table 1

The Distribution of Research Articles across Publication Years (2014-January & 2018-June)

Year	EFL In-service Context		EFL Pre-service Context		Total
	Foreign*	Turkish	Foreign	Turkish	
2018	2	2	-	1	5
2017	5	2	-	-	7
2016	1	2	1	-	4
2015	2	-	-	1	3
2014	1	-	1	-	2
Total	11	6	2	2	21
	17		4		

*Foreign refers to the countries which have an EFL context other than Turkey such as China, Colombia and the like in order to show the comparison.

As for the features of the articles, they were outlined in Table 2. When this table is scrutinized, it can be seen that most researchers have preferred qualitative research designs ($f=12$). It can be assumed that more different research designs may be complementary to enlighten the issue of LAL from different perspectives. In addition, more different research tools can be utilized because interviews and questionnaires/surveys were dominant.

Table 2

The Features of LAL Research Articles between 2014-January and 2018-June

	Author(s)/Year	Type	Context	Participants/Instruments
1	Giraldo Aristizabal (2018)	action research	in-service EFL (Colombia)	60 teachers questionnaire, interview, document
2	Yan, Zhang & Fan (2018)	qualitative	in-service EFL (China)	3 EFL teachers interview

3	H. Xu (2017)	longitudinal (3 years)	in-service EFL- novice (China)	4 novice teachers observation, interview, field notes, journal writing
4	Tsagari & Vogt (2017)	qualitative	in-service EFL (Cyprus, Germany, Greece)	63 teachers interview
5	Y. Xu & Brown (2017)	quantitative	in-service EFL (China)	891 teachers questionnaire, survey
6	Sellan (2017)	qualitative	in-service EFL (Singapore)	8 teachers documents, interview, stimulated recall, observation
7	Baker & Riches (2018)	qualitative	in-service EFL training (Haiti)	120 teachers feedback, critiques, survey, interview
8	Y. Xu (2016)	qualitative	in-service EFL (China)	20 teachers lesson plans, interviews
9	Hakim (2015)	quantitative	in-service EFL (Saudi Arabia)	30 instructors survey
10	Jannati (2015)	qualitative	in-service EFL (Iran)	18 instructors interview
11	Vogt & Tsagari (2014)	mixed	in-service EFL (EU countries)	7 countries questionnaire: 853 interview: 63
12	Yastibas & Takkac (2018a)	qualitative	in-service EFL (Turkey)	8 instructors think-aloud protocols
13	Yastibas & Takkac (2018b)	qualitative	in-service EFL (Turkey)	8 instructors interview, focus group
14	Mede & Atay (2017)	mixed	in-service EFL (Turkey)	350 instructors questionnaire, focus group
15	Oz & Atay (2017)	qualitative	in-service EFL (Turkey)	12 instructors interview
16	Tuzcu-Eken (2016)	qualitative	in-service EFL (Turkey)	5 teachers & 15 students interview
17	Buyukkarci (2016)	quantitative	in-service EFL (Turkey)	32 teachers survey
18	Viengasang (2016)	mixed	pre-service ELT (Thailand)	46 undergraduates at practicum survey, interview
19	Lam (2014)	qualitative	pre-service ELT (Hong Kong)	5 teacher education institutions program and government documents, interview, student assessment tasks, teaching evaluation

20	Komur (2018)	qualitative	pre-service ELT (Turkey)	49 undergraduates open-ended questions
21	Hatipoglu (2015)	mixed	pre-service ELT (Turkey)	124 undergraduates survey, interview

As a conclusion, there have been noteworthy attempts in the literature to investigate the concept of LAL within foreign language education contexts.

Findings

Empirical research articles about EFL teachers' LAL were analyzed in terms of their teaching contexts (in-service vs pre-service). It is because teaching context is considered to affect and shape all the assessment practices and knowledge of language teachers. The concept of context includes learners' characteristics and needs, teaching/learning goals, teachers' experience in assessment, teachers' own beliefs, resources, training phase and the like within foreign language education. Therefore, this review has divided findings into two sections: In-service context and pre-service context.

In-service EFL Context of LAL Research

This section of the present review is mainly about the recent studies carried out within in-service EFL context regarding LAL. Specifically, it illustrates what researchers have investigated in terms of EFL teachers' language assessment, whether EFL teachers who are working in schools or at universities are language assessment literate, what kind of variables have an impact on their LAL, and whether they need training on language assessment or not.

To begin with, Giraldo Aristizabal (2018) investigated the assessment beliefs and practices of English language instructors working at a language institute in Colombia while designing an achievement test. He carried out his research as a part of action research which aimed to develop LAL of in-service EFL teachers. 60 EFL teachers participated in the study. Since that part of the research was qualitative-oriented, questionnaires, interviews and documents were the primary data collection tools. The researcher used a priori coding system in order to analyze the data. The analysis yielded that the participant teachers believed tests should have basic qualities such as validity and reliability while they reported that not all qualities were reflected in the practices. They also indicated that designing a test as a part of their training was useful for their professional development.

Yan, Zhang, and Fan (2018), however, focused on contextual and experiential factors in the development of LAL. That is, they designed their study in order to evaluate how such factors interact with each other to enhance LAL. 3 EFL teachers from middle school in China were interviewed with five questions regarding their experiences in assessment practices in terms of their working context. The analysis was conducted through an inductive approach with five categories as assessment context, assessment training experience, assessment practice, assessment knowledge and assessment training needs. The findings revealed a separate LAL profile of teachers in which the effect of context was found. The participants also needed more training in

assessment practice rather than assessment knowledge. In this sense, the results showed that their theoretical knowledge was mostly formed through their practices and working contexts. Thus, teachers had fine opinions about assessment issues due to their practices.

In a similar vein, Xu (2017) explored the development of LAL of novice EFL teachers. He carried out a three-year-longitudinal study with 4 EFL teachers in China. He collected qualitative data with the help of different techniques such as classroom observation, interview, journal entries and fieldnotes. As a result of his analysis of 3-year-development, he found three stages: The first stage was learning and applying practical assessment techniques; the second stage was connecting assessment to teaching goals; the third stage was the transition from planned assessment to more improvised formative assessment. Those stages reflected the participants' progress in their language assessment practices. The researcher also emphasized that improvised assessment may be the sign of LAL development.

Tsagari and Vogt (2017) concentrated on the qualitative part of their study conducted across seven European countries in 2014 in order to discuss perceived levels of LAL as well as training needs of in-service EFL teachers. However, the researchers took only three countries; namely, Cyprus, Germany, and Greece, for that part of the study. Thus, they conducted interviews with 63 EFL teachers from primary or secondary school context. Teaching experience of the participants ranged from two to thirty-four years. The interview findings indicated that nearly all the teachers still utilized pen-and-paper tests and traditional assessment tasks but did not use standardized tests; they preferred classroom testing. Furthermore, teachers reported that they knew alternative assessment tools but did not use in their classroom practices. The participants also mentioned they received language assessment training but found it not very useful because their LAL seemed to improve during working professionally. However, they were not successful enough to determine what they need as assessment training or what they need to enhance their LAL levels. So, the researchers argued that the sample was not at a sufficient level in terms of LAL, which indicated that language assessment training in teacher education was not adequate. They also commented that language assessment practices of EFL teachers were not much developed.

Likewise, Xu and Brown (2017) intended to investigate EFL teachers' levels of LAL and the impact of demographic variables on their assessment performances. Therefore, they applied an adapted version of teacher assessment literacy questionnaire surveys consisting of assessment scenarios and also items related to demographic information. The sample was made up of 891 EFL teachers working at universities from seven regions of China. The results of the surveys revealed that demographic variables did not influence the levels of LAL, i.e. teaching experience did not affect assessment decisions or performances much. Besides, the levels of LAL were at the moderate level. Therefore, the researchers recommended that the concept of LAL and its measurement should be contextually grounded so that it can reflect the assessment contexts where teachers work.

Unlike other studies, Sellan (2017) focused on a new language assessment paradigm introduced in Singapore as a part of educational policy. That was Integrated Program which aimed to stimulate teacher-based classroom assessment practices to enhance language learning while delimitating traditional exams. Thus, the study was

based on EFL teachers' formative assessment and how such assessment was helpful in language learning. 8 high school EFL teachers were interviewed three times: in the beginning, in the middle, in the end. Lesson observations, policy documents and paper marking practices were also evaluated. The analysis of all the data collected yielded that EFL teachers not only improved in terms of language related issues such as knowledge, genre, culture and authenticity but also performed better assessment techniques in line with their expanded language improvement. In this way, teachers stated that they could more easily and successfully respond to the needs of language learners.

From the perspective of in-service training, Baker and Riches (2018) designed a one-week-workshop so as to improve the participants' levels of LAL in Haiti. Hence, they examined the development of LAL from the views of two stakeholders: EFL teachers and facilitators. 120 EFL teachers were working in high schools and did not take an assessment training before. During language assessment workshop, the researchers collected teachers' feedback on draft exams as well as their critiques of national exams. They also conducted surveys and interviews with the sample. They found that both types of participants developed positively in language assessment. However, they had challenges in decision-making phase while collaboratively studied with facilitators, but the researchers compromised with teachers more easily, which showed that teachers were more open to development.

On the other hand, Xu (2016) was motivated by the problem that was not much explored: how teachers planned their assessment practices. The findings for this problem would also show the development level of LAL. So, she run qualitative analyses of EFL teachers' lesson plans and additionally, conducted interviews with 20 university teachers in China. As a consequence of her analysis, it was found that teachers obviously added assessment components such as objectives in their plans but did not match learning objectives with their assessment practices. Teachers also did not utilize rubrics for performance assessment. The researcher concluded that EFL teachers' assessment was much more intuitive and governed by their instruction rather than clearly planned.

Considering the relationship between experience and assessment, Hakim (2015) explored the perceptions of EFL teachers in line with their experiences in assessment in order to reveal the level of awareness and ideology regarding LAL. She developed a questionnaire related to language assessment perceptions and 30 language instructors working in English Language Institute in Saudi Arabia responded. The questionnaire responses showed that the participants had the knowledge of assessment techniques and their use but were not able to put that knowledge into practice. Therefore, teaching experience did not affect their assessment practices, but affected only understanding of assessment components or testing plan.

In the same vein, Jannati (2015) focused on the perceptions and practices of in-service English language teachers about assessment. Her study included 18 English language instructors who were divided into three groups with regard to their teaching experience from low to highly experienced. The findings from interviews indicated that the instructors were familiar with general concepts and terminologies of assessment. Nevertheless, teaching experience did not influence instructors' perceptions of assessment. So, though they had the knowledge of assessment, which was their assessment literacy, they were not able to use that literacy in their teaching practices.

Pointing to training needs, Vogt and Tzagari (2014) investigated the levels of EFL teachers' LAL and whether they needed training in assessment across seven European countries (Cyprus, Macedonia, Germany, Greece, Italy, Poland, Turkey) by means of questionnaires. As mentioned in the previous parts, the researchers did the same study but divided it into two parts as qualitative in 2017 and quantitative in 2014. Consequently, that study was the quantitative part of their research. The researchers found that teachers needed training in language assessment though they were working professionally. It is because local needs of educational contexts were different, and their knowledge of assessment could not be adequate to assess language learners in practice. Therefore, the researchers highlighted that pre-service teacher education is important in providing a good model of language assessment training.

Considering Turkish in-service context, Yastibas and Takkac (2018a) were interested in how EFL instructors decided what to assess and how to assess. In this way, they aimed to show how teachers' LAL would be reflected in their practices. Therefore, they made 8 EFL instructors from a Turkish university prepare and design language tests for different language skills and areas. They utilized think-aloud protocols with which they tried to illustrate the process of planning assessment. The findings indicated that the participants prepared language tests according to the progress of their students and the material used as coursebook. Therefore, all the assessment was structured in line with students and coursebook. The sample also reported that language tests had a positive washback on students learning.

In another study, Yastibas and Takkac (2018b) investigated how EFL teachers developed their LAL while working as language instructors at university in Turkey. In their study, 8 EFL instructors were individually interviewed and also were included in focus group discussion about their language assessment conceptions and experiences. The researchers shaped the findings in three dimensions: previous assessment experience, assessment training and self-improvement. It was found that teachers' previous assessment experiences differed from each other, and thus, they held different beliefs and conceptions about language assessment, which reflected their assessment practices. Moreover, teachers' pre-service assessment training was much more influential in their development of LAL. Among those three dimensions, the most prominent one was found as self-improvement because the participants reported that while they had the chance to practice their assessment knowledge by preparing and administering assessment procedures, they gained experience, which led to an improvement in their LAL level. Hence, the researchers highlighted that language assessment experience was an important factor in the improvement of LAL.

Mede and Atay (2017) were also concerned with the levels of LAL of EFL teachers who were working at universities in Turkey. They designed their study to find out training needs in language assessment practices. They collected relevant data by means of a questionnaire and focus group interviews. 350 EFL instructors with more than five years of teaching experience took part in the study. The findings indicated that the whole sample needed training even in basic assessment techniques. The sample stated that they were not able to prepare good language tests and provide feedback according to the test results. They also mentioned that the construction of skill-based language tests was difficult for them; they were only good at preparing vocabulary and grammar tests. So, the researchers pointed out that the assessment training the sample

received was not sufficient because it did not improve language assessment practices except for grammar and vocabulary testing.

In addition, Oz and Atay (2017) focused on in-class language assessment perceptions of Turkish EFL instructors. They investigated teachers' reflections, knowledge, practices and experiences with respect to language assessment through their perceptions. 12 EFL instructors were interviewed and they expressed that they had much knowledge about language assessment. But it was problematic when teachers tried to apply that knowledge into their assessment skills, which indicated that LAL could not be reflected in classroom assessment well. Their teaching experience also was not very helpful in improving LAL in practice.

From a different point of view, Tuzcu-Eken (2016) addressed the problems of testing and obstacles encountered within vocational school context in the implementation of assessment. To investigate those issues, she conducted interviews with 5 Turkish EFL teachers about their views on testing and what kind of problems they came across while assessing language learners. She also put emphasis on students' opinions in the process; therefore, she carried out focus group interviews with 15 students. While students were found to have lack of motivation since their main goal was to pass the course, teachers reported that they tried their best to employ different language assessment techniques despite certain barriers. The problems mentioned comprised of physical conditions such as lack of technological tools, crowded classes and not administering listening and speaking assessments properly due to class size. Though teachers believed they were somewhat language assessment literate, they could not reflect their literacy into their contexts, which shows that external factors hindered applying appropriate and comprehensive language assessments.

Buyukkarci (2016), however, carried out his research in order to find out the levels of LAL of Turkish EFL teachers working in primary, secondary and tertiary education. He also explored whether teaching experience and post-graduate education led to a difference in teachers' LAL. 32 EFL teachers responded to a survey of assessment literacy. The results yielded low levels of LAL, which indicated that teachers' LAL did not improve much. Furthermore, both teaching experience and doing post-graduate education did not contribute to the development of the participants' LAL. All those results produced the interpretations that pre-service training, in-service teaching experience and post-graduate training were not really helpful in enhancing LAL; so, teachers needed satisfactory language assessment training.

Pre-service EFL Context of LAL Research

This section of the current review is based on the most up-to-date language assessment studies conducted with pre-service EFL teachers who would be teachers at the end of their teacher training. Essentially, it depicts language assessment training of prospective EFL teachers, to what extent they are language assessment literate, what kind of variables have an impact on their LAL during their teacher education, and whether the training was useful or not. By addressing those questions, the researchers have sought the ways to improve LAL of pre-service teachers. However, it should be noted that in recent years, there are not many studies interested in pre-service context compared to in-service context.

To start with, Viengsang (2016) carried out her study in order to investigate how assessment literacy was understood, the effect of previous training, and practices in the practicum, the needs and problems regarding assessment. She was concerned with pre-service language assessment training context, and hence, 46 EFL teacher candidates doing their teaching practice in Thailand took part in the study. The sample firstly responded to a questionnaire and then, 5 of them were interviewed. The results showed that though the sample was trained, and had the knowledge of language assessment, they were not able to reflect that knowledge to their classrooms. Hence, a gap was revealed between theoretical side and practical side of language assessment course taken in training process.

Differently, Lam (2014) designed a comprehensive research which aimed to explore the overall language assessment training programs offered in five Hong Kong English language teacher education institutions. More specifically, whether such programs were facilitative or debilitating in the development of teachers' language assessment literacy was investigated. A survey of the programs and government documents, interviews with selected pre-service teachers and instructors, the evaluation of students' assessment tasks, and the evaluation of assessment courses were utilized to investigate the effect of the programs. Five main themes were reported: the assessment courses in the programs were not effective in enhancing LAL; the application of LAL in real classroom contexts was limited; there was a lack of training in administering assessments of both classroom-based and large-scale; the experiences of the participants in the course were varied, either being positive or negative; the perceptions related assessment was examination-oriented due to their exam-oriented culture. Therefore, in general, the assessment courses were not as effective as expected in developing LAL of pre-service language teachers.

In terms of Turkish pre-service context, Komur (2018) based his research on the awareness and readiness in terms of language assessment. So, he explored Turkish pre-service EFL teachers' knowledge about testing and also questioned whether the content of their assessment training course was sufficient or not in responding to their language assessment needs. 49 senior undergraduates who were prospective EFL teachers answered four open-ended questions. The analysis of their responses yielded that most of the participants were aware of recent developments in ELT field but needed more training on practical techniques of assessment corresponding to the latest innovation in ELT. Therefore, they perceived themselves not qualified enough to conduct an appropriate language assessment despite their expanded knowledge of theories of testing.

Similarly, Hatipoglu (2015) carried out a three-year-long research in order to explore the needs of pre-service EFL students regarding language assessment and testing. She collected needs analysis surveys and interviews from 124 undergraduate students of ELT. The findings indicated that local needs, contexts, and students' previous experiences with assessment had an impact on students' beliefs about English language testing and evaluation training, and one single course was not enough because the students had limited knowledge in language testing and assessment. Therefore, she concluded that the content of the course required changes in order for better training in language assessment.

Discussion and Conclusion

In this section, the interpretive systematic review of the literature and current research studies carried out about LAL have been summarized, synthesized, and also criticized in terms of missing points. Moreover, implications for further practice and research; that is, future directions, have been recommended. Hence, the past, present and future of LAL concept have been illustrated.

The effects of LAL have been recognized when the importance of the relationship between teaching and language assessment is acknowledged. It is because assessment is one of the helpful ways in evaluating the quality of education. However, the concept of LAL research can be thought new compared to AL research (Zolfaghari & Ahmadi, 2016). As a consequence of LAL research, there has been an increasing interest to define what LAL is, to show how it is developed, to find what kind of features it has and to illustrate how it is implemented in classrooms in order to emphasize its role in foreign language education (Csépes, 2014).

As far as the conceptual side of LAL is taken into consideration, this concept has been defined, conceptualized and characterized. Terminologically, LAL refers to the familiarity with assessment procedures and effectively performing them in teaching contexts. The definitions of LAL provided up to this time have had common aspects such as teachers' knowledge, skills, principles, contexts, procedures and understanding of language assessment within the perspective of language teaching education (e.g. Davies, 2008; Fulcher, 2012; Inbar-Lourie, 2008). Nonetheless, it is clear that more research into this multi-layered concept is needed because only one stakeholder (teacher) has been studied mostly, and less attention has been given to the importance of different language contexts such as in-service and pre-service teachers, other stakeholders or cultural issues as well as the ethics in language assessment. Therefore, by taking into account these features, other definitions of LAL might be formed covering all the possible layers of language assessment or different definitions according to different stakeholders apart from Taylor's (2013) model and contexts for a better understanding of this concept. When components or dimensions of LAL are considered, there have been valuable attempts with regard to these in the literature. For instance, Davies (2008), Fulcher (2012), Inbar-Lourie (2008), and Taylor (2013) have portrayed the dimensions of LAL according to their definitions or understandings. It can be deduced that LAL possesses more or less the same components or dimensions with different classifications such as knowledge, process, skills, principles, abilities and contexts. However, Davies' (2008) classification of LAL as knowledge, skills and principles has been appreciated much more in educational contexts. On the other hand, the challenge here is to show how LAL components can be operationalized in language teaching practices. For example, Fulcher (2012) has tried to construct a working definition of LAL by doing research with in-service foreign language teachers. Yet, it is better that theory and practice of LAL should be thought together for different contexts as well as for different stakeholders. In addition to these, the researchers have characterized what kind of features make language teachers assessment literate. Most of them have indicated common characteristics such as being aware of what sound and unsound assessment is, knowledge of assessment theory and pedagogical ways to perform it, ability to match learning outcomes with assessment processes, applying appropriate language assessment tools and so on (e.g. Huang & He, 2016; Inbar-Lourie,

2013b; Khadijeh & Amir, 2015; Rogier, 2014). On the other hand, while some researchers have stated the characteristics of LAL could be valid for all the stakeholders concerned in assessment, some others have stressed that for each stakeholder or academic discipline, characteristics may differ and thus, such characteristics can be taken into consideration. For aforementioned reasons, research into this gap can be useful for educational concerns to develop a deeper understanding of LAL. Besides, the feature of ethics has not been highlighted much; therefore, the ethical side of LAL can be studied in order to contribute to the conceptualization of LAL.

In terms of the research executed about LAL, it is apparent that there are several reviews, theoretical or conceptual studies in the literature. But there is not much exploration about how LAL can be operationalized in foreign language teaching/learning contexts or how it can differ between pre-service and in-service EFL teachers. Therefore, much empirical research is needed to enlighten the concept of LAL as in Fulcher's (2012) study in which both theory and practice of LAL have been investigated. In other words, doing research to construct theory or conceptualization is required instead of discussing LAL in its own. In addition, the outline of studies has yielded that the number of qualitative studies is greater than the other research designs, and interviews as the data collection instrument are dominant. For that reason, various research designs and instruments can be employed so as to clarify and interpret the meaning and implementation of LAL better.

As previously stated, this review covers the empirical research of LAL performed in in-service and pre-service foreign language education contexts between the years of 2014-January and 2018-June. From this perspective, it can be seen that there are more studies with in-service EFL teachers than pre-service EFL teachers.

With regard to in-service EFL teacher context, most of the studies have concluded that the levels of EFL teachers' LAL were low and thus, they needed language assessment training (e.g. Buyukkarci, 2016; Mede & Atay, 2017; Tsagari & Vogt, 2017; Vogt & Tsagari, 2014; Yan, Zhang & Fan, 2018). This training can be a complementary in-service training. From this point of view, Giraldo Aristizabal (2018) and Baker and Riches (2018) found positive effects of workshops regarding language assessment on teachers' development of LAL. So, other in-service training sessions may be designed if teacher education programs fall behind in this respect (Popham, 2009). In a different way, the needs of training are considered as a sign of inadequate pre-service teacher education. Therefore, remedial actions can be taken to improve the content of language assessment courses in foreign language teacher education programs. If pre-service EFL teachers took a good training on language assessment, they would be more successful in their assessment practices as in Yastibas and Takkac's (2018b) study in which previous assessment courses affected the development of LAL in a positive way. Even, if this kind of strategy were taken as a policy like in Sellan's (2017) study, training programs would produce more capable and devoted foreign language teachers in terms of language assessment. In addition, some studies that have investigated the effect of demographic variables and contextual issues have found that the context where teachers work affects their assessment practices and sometimes, EFL teachers could not respond to the needs of that context (e.g. Tuzcu-Eken, 2016; Vogt & Tsagari, 2014; Yan, Zhang & Fan, 2018). On the other hand, most of demographic variables have not influenced LAL. Specifically, researchers have focused on teaching experience but have

not found a positive or facilitative effect in improving LAL of EFL teachers (e.g. Buyukkarci, 2016; Hakim, 2015; Jannati, 2015; Oz & Atay, 2017; Xu & Brown, 2017). It can be assumed that training and context are crucial factors whereas experience does not make a difference in LAL enhancement. Finally, nearly all the studies with in-service EFL teachers reviewed have revealed that teachers have some knowledge of language assessment, but they are not able to put that knowledge into practice. Accordingly, the practical side of language assessment should be highlighted.

Considering pre-service EFL teacher context, as noted before, there are not many empirical studies in recent years. However, up-to-date research into this context has shown that there is a gap between theory and practice of language assessment as discussed by in-service EFL teachers. In other words, the finding of the research in the last five years has indicated that prospective EFL teachers complained about the emphasis put more on theory rather than practice (e.g. Komur, 2018; Viengsang, 2016). Thus, a balance between theory and practice is required for better language assessment training. For this goal, the training of LAL can encompass “an appropriate balance of technical know-how, practical skills, theoretical knowledge, and understanding of principles, but all firmly contextualized within a sound understanding of the role and function of assessment within education and society” (Taylor, 2009, p. 27). In addition, because of this kind of imbalance, EFL teacher candidates were found to have lower levels of LAL (e.g. Komur, 2018). Likewise, the training courses on language assessment in foreign language education have appeared to be ineffective in supporting the development of LAL and the content of such courses have fallen short in equipping their undergraduate students with necessary knowledge and skills required for good practices of language assessment (e.g. Hatipoglu, 2015; Lam, 2014). All of these studies have concluded that language assessment training programs in foreign language education should be improved in order to provide a better education. It is because to develop LAL of teacher candidates is believed to be a necessary component of foreign language teacher training program (DeLuca & Klinger, 2010; Herrera & Macias, 2015). For instance, the content of the programs may be revised and renewed according to the needs of their attenders as well as the developments and changes in foreign language education and assessment field. One of the points also attracting attention is prospective EFL teachers’ inability to interpret assessment results (Lam, 2014). Therefore, the training courses should equally give importance to all the required characteristics of language assessment literate teachers. Besides, to focus on the needs of LAL training might be useful to graduate qualified EFL teachers. With respect to this idea, it should be emphasized that pre-service language teacher training is a critical phase in developing teacher competencies such as language assessment. Thus, more research can be carried out to show what may be done to improve LAL of teacher candidates and how foreign language education programs might be shaped in line with such directions.

After all, this systematic review is an attempt to demonstrate the current place of LAL, its importance in the literature, and strengths and weaknesses of studies conducted in recent five years in terms of foreign language education context. It has also provided some implications for practice and recommendations for further research. It is hoped that this review has contributed to the LAL research field by discussing, summarizing, synthesizing and suggesting. Furthermore, it is believed that this review reveals new

perspectives regarding LAL. Still, researchers must carry on investigating LAL in order to shed light on its hidden points.

References

- *Baker, B. A., & Riches, C. (2018). The development of EFL examinations in Haiti: Collaboration and language assessment literacy development. *Language Testing*, 35(4), 557-581. doi:10.1177/0265532217716732
- Bayat, K., & Rezaei, A. (2015). Importance of teachers' assessment literacy. *International Journal of English Language Education*, 3(1), 139-146.
- Berry, V., & O'Sullivan, B. (2016). *Assessment literacy for language teachers*. Retrieved from <http://tea.iatefl.org/wp-content/uploads/2015/10/Vivien-Berry-and-Barry-OSullivan.pdf>
- *Buyukkarci, K. (2016). Identifying the areas for English language teacher development: A study of assessment literacy. *Pegem Egitim ve Ogretim Dergisi [Pegem Journal of Education and Instruction]*, 6(3), 333-346. doi:10.14527/pegegog.2016.017
- Cooper, H., Hedges, L. V., & Valentine, J. C. (2009). *The handbook of research synthesis and meta-analysis* (2nd ed.). NY, USA: Russell Sage Foundation.
- Csépes, I. (2014). Language assessment literacy in English teacher training programmes in Hungary. In J. Hovarth & P. Medgyes, (Eds.), *Studies in Honour of Marianne Nikolov* (pp. 399-411). Pécs: Lingua Franca Csoport.
- Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing*, 25(3), 327-347. doi:10.1177/0265532208090156
- DeLuca, C., & Klinger, D. A. (2010). Assessment literacy development: Identifying gaps in teacher candidates' learning. *Assessment in Education: Principles, Policy & Practice*, 17(4), 419-438. doi:10.1080/0969594X.2010.516643
- Djoub, Z. (2017). Assessment literacy: Beyond teacher practice. In R. Al-Mahrooqi, C. Coombe, F. Al-Maamari & V. Thakur (Eds.), *Revisiting EFL assessment* (pp. 9-27). Switzerland: Springer International Publishing.
- *Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113-132. doi:10.1080/15434303.2011.642041
- Giraldo, F. (2018). Language assessment literacy: Implications for language teachers. *Profile: Issues in Teachers' Professional Development*, 20(1), 179-195. doi:10.15446/profile.v20n1.62089
- *Giraldo Aristizabal, F. G. (2018). A Diagnostic study on teachers' beliefs and practices in foreign language assessment. *Íkala: Revista de Lenguaje y Cultura*, 23(1), 25-44. doi:10.17533/udea.ikala.v23n01a04
- Gotch, C. M., & French, B. F. (2014). A systematic review of assessment literacy measures. *Educational Measurement: Issues and Practice*, 33(2), 14-18.
- *Hakim, B. (2015). English language teachers' ideology of ELT assessment literacy. *International Journal of Education & Literacy Studies*, 3(4), 42-48. doi:10.7575/aiac.ijels.v.3n.4p.42
- *Hatipoglu, C. (2015). English language testing and evaluation (ELTE) training in Turkey: expectations and needs of pre-service English language teachers. *ELT Research Journal*, 4(2), 111-128.
- Herrera, L., & Macias, D. (2015). A call for language assessment literacy in the education and development of teachers of English as a foreign language.

- Colombian Applied Linguistics Journal*, 17(2), 302-312.
doi:10.14483/udistrital.jour.calj.2015.2.a09
- Huang, J., & He, Z. (2016). Exploring assessment literacy. *Higher Education of Social Science*, 11(2), 18-27. doi:10.3968/8727
- Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*, 25(3), 385-402. doi:10.1177/0265532208090158
- Inbar-Lourie, O. (2013a). Language assessment literacy. In C. A. Chapella (Ed.), *The Encyclopedia of Applied Linguistics* (pp. 1-9). Blackwell Publishing Ltd. doi:10.1002/9781405198431.wbeal0605
- Inbar-Lourie, O. (2013b, November). *Language assessment literacy: What are the ingredients?* Paper presented at the 4th EALTA-CBLA SIG Symposium, Nicosia, Cyprus. Retrieved from http://www.ealta.eu.org/events/CBLA-cyprus2013/Lectures_workshops/O_Inbar-Lourie%20-%20plenary%20-1.pdf
- Inbar-Lourie, O. (2017). Language assessment literacy. In E. Shohamy, L. Or & S. May (Eds.), *Language Testing and Assessment* (pp. 257-270). Springer International Publishing AG. doi:10.1007/978-3-319-02261-1_19
- *Jannati, S. (2015). ELT teachers' language assessment literacy: Perceptions and practices. *The International Journal of Research in Teacher Education*, 6(2), 26-37.
- Khadijeh, B., & Amir, R. (2015). Importance of teachers' assessment literacy. *International Journal of English Language Education*, 3(1), 139-146. doi:10.5296/ijele.v3i1.6887
- *Komur, S. (2018). Preservice English teachers' assessment awareness: Level of readiness for classroom practice. *The Journal of Language Teaching and Learning*, 8(1), 109-121.
- *Lam, R. (2014). Language assessment training in Hong Kong: Implications for language assessment literacy. *Language Testing*, 32(2), 169-197. doi:10.1177/0265532214554321
- *Mede, E., & Atay, D. (2017). English language teachers' assessment literacy: The Turkish context. *Dil Dergisi-Ankara Universitesi TOMER [Language Journal-Ankara University TOMER]*, 168(1), 43-60.
- Mertler, C. A. (2003, October). *Pre-service versus In-service teachers' assessment literacy: Does classroom experience make a difference?* Paper presented at the Annual Meeting of the Mid-Western Educational Research Association, Columbus, OH, USA. Retrieved from the ERIC database (ED482277).
- *Oz, S., & Atay, D. (2017). Turkish EFL instructors' in-class language assessment literacy: Perceptions and practices. *ELT Research Journal*, 6(1), 25-44.
- Petticrew, M., & Roberts, H (2006). *Systematic reviews in the social sciences: A practical guide*. Oxford, UK: Blackwell Publishing.
- Pill, J., & Harding, L. (2013). Defining the language assessment literacy gap: Evidence from a parliamentary inquiry. *Language Testing*, 30(3), 381-402. doi:10.1177/0265532213480337

- Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory into Practice*, 48(4), 4-11. doi:10.1080/00405840802577536
- Rea-Dickins, P. (2004). Understanding teachers as agents of assessment. *Language Testing*, 21(3), 249-258. doi:10.1191=0265532204lt283ed
- Rogier, D. (2014). Assessment literacy: Building a base for better teaching and learning. *English Language Teaching Forum*, 52(3), 2-13.
- Scarino, A. (2013). Language assessment literacy as self-awareness: Understanding the role of interpretation in assessment and in teacher learning. *Language Testing*, 30(3), 309-327. doi:10.1177/0265532213480128
- *Sellan, R. (2017). Developing assessment literacy in Singapore: How teachers broaden English language learning by expanding assessment constructs. *Papers in Language Testing and Assessment*, 6(1), 64-87.
- Stabler-Havener, M. L. (2018). Defining, conceptualizing, problematizing, and assessing language teacher assessment literacy. *Teachers College, Columbia University Working Papers in Applied Linguistics & TESOL*, 18(1), 1-22.
- Stiggins, R. J. (1995). Assessment literacy for the 21st century. *Phi Delta Kappan*, 77(3), 238-245.
- Taylor, L. (2009). Developing assessment literacy. *Annual Review of Applied Linguistics*, 29, 21-36. doi:10.1017/S0267190509090035
- Taylor, L. (2013). Communicating the theory, practice and principles of language testing to test stakeholders: Some reflections. *Language Testing*, 30(3), 403-412. doi:10.1177/0265532213480338
- *Tzagari, D., & Vogt, T. (2017). Assessment literacy of foreign language teachers around Europe: Research, challenges and future prospects. *Papers in Language Testing and Assessment*, 6(1), 41-63.
- *Tuzcu-Eken, D. (2016). Testing in English classes in vocational schools of higher education. *Electronic Journal of Vocational Colleges*, 6(1), 38-47.
- *Viengsang, R. (2016). Exploring pre-service English teachers' language assessment literacy. *Modern Journal of Language Teaching Methods (MJLTM)*, 6(5), 432-442.
- *Vogt, K., & Tzagari, D. (2014). Assessment literacy of foreign language teachers: Findings of a European study. *Language Assessment Quarterly*, 11(4), 374-402. doi:10.1080/15434303.2014.960046
- Wach, A. (2012). Classroom-based language efficiency assessment: A challenge for EFL teachers. *Glottodidactica*, 39(1), 81-92.
- *Xu, H. (2017). Exploring novice EFL teachers' classroom assessment literacy development: A three-year longitudinal study. *Asia-Pacific Educational Research*, 26(3-4), 219-226. doi:10.1007/s40299-017-0342-5
- *Xu, Y. (2016). Assessment planning within the context of university English language teaching (ELT) in China: Implications for teacher assessment literacy. *Australian Review of Applied Linguistics*, 39(3), 233-254. doi:10.1075/aral.39.3.o2xu
- *Xu, Y., & Brown, G. T. L. (2017). University English teacher assessment literacy: A survey-test report from China. *Papers in Language Testing and Assessment*, 6(1), 133-158.

- *Yan, X., Zhang, C., & Fan, J. J. (2018). “Assessment knowledge is important but ...”: How contextual and experiential factors mediate assessment practice and training needs of language teachers. *System*, 74, 158-168. doi:10.1016/j.system.2018.03.003
- *Yastibas, A. E., & Takkac, M. (2018a). Understanding language assessment literacy: Developing language assessments. *Journal of Language and Linguistic Studies*, 14(1), 178-193.
- *Yastibas, A. E., & Takkac, M. (2018b). Understanding the development of language assessment literacy. *Bingol University Journal of Social Sciences Institute*, 8(15), 89-106. doi:10.29029/busbed.364195
- Zolfaghari, F., & Ahmadi, A. (2016). Assessment literacy components across subject matters. *Cogent Education*, 3, 1-16. doi:10.1080/2331186X.1252561

Note: References marked with an asterisk indicate the studies in this review.



This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0). For further information, you can refer to <https://creativecommons.org/licenses/by-nc-sa/4.0/>