

Developing EFL Teachers' Language Assessment Literacy: A Systematic Literature Review on Teacher Training Programs

August 2024 – Volume 28, Number 2

<https://doi.org/10.55593/ej.28110a8>

Indah Puspawati

Universitas Negeri Malang, Indonesia;
Universitas Muhammadiyah Yogyakarta, Indonesia
<indahpuspawati@umy.ac.id>

Maharani Khansa

Universitas Negeri Malang, Indonesia
<maharani.khansa.2202219@students.um.ac.id>

Utami Widiati

Universitas Negeri Malang, Indonesia
<utami.widiati.fs@um.ac.id>

Abstract

Teachers' language assessment literacy (LAL) refers to teachers' ability to perform assessment-related tasks. Despite the significant role of language assessment literacy in English language teaching and learning, many teachers reported to have an inadequate ability to perform assessment-related tasks because of the lack of assessment training programs that prepared them for the tasks. This study aimed at investigating the existing assessment training programs to develop EFL teachers' LAL. A systematic literature review was conducted to observe trends in how assessment training programs were carried out in the past decades. PRISMA model was used for identifying, screening, and selecting. Twenty-eight studies published between 2012 and 2022 were reviewed. The results reveal that the assessment training models comprised theoretical and practical components of language assessments. The results also show that the methods for delivering the content are collaborative and reflective practices. The findings of this study should be considered in the design of EFL teacher training programs on assessment.

Keywords: Language assessment literacy, LAL, EFL teachers, Assessment training, Systematic literature review

Teachers should have the ability to conduct assessments in teaching and learning. For language teachers, the term 'Language Assessment Literacy' (LAL) refers to language teachers' ability to design, administer, and interpret the result of assessments (Davies, 2008; Fulcher, 2012; Giraldo, 2020; Inbar-Lourie, 2008; Sevimeel-Sahin & Subasi, 2019). Previous research revealed

that teachers' LAL impacted their teaching (Ahmadi et al., 2022; Mellati & Khademi, 2018; Sultana, 2019). Sultana's (2019) research reported that teachers with adequate LAL knowledge delivered better language teaching. Mellati and Khademi's (2018) research also suggests that teachers with adequate LAL performed better in their teaching because they could consistently evaluate their teaching through the results of assessments. Other studies indicated how teachers with low or insufficient LAL failed to respond well to education reforms in their countries and could not adapt their teaching and assessment practices as mandated by the reformed curriculum (Firoozi et al., 2019; Razavipour & Rezagah, 2018). Thus, teachers' ability to perform assessments is equally essential to teachers' ability to teach.

Despite the significant role of LAL, most studies on teacher training and professional development programs focused on teaching methodologies (Sayyadi, 2022; Sulaiman et al., 2021). Teachers claimed that they were more knowledgeable and confident in their teaching skills than their assessment skills (Ahmadi & Ketabi, 2020; Bahtiar & Purnawarman, 2020; Fitriyah et al., 2022; Latif, 2021). Moreover, many studies on teachers' LAL indicated that teachers have inadequate LAL to perform assessment-related tasks (Bøhn & Tsagari, 2021; Firoozi et al., 2019; Liu & Li, 2020; Ölmezer-Öztürk & Aydin, 2018; Razavipour & Rezagah, 2018; Sultana, 2019). Even if the teachers were reported to have sufficient or excellent knowledge of LAL, they failed to transform their knowledge into assessment practices (Aria et al., 2021; Levi & Inbar-Lourie, 2020; Öz & Attay, 2017; Zulaiha et al., 2020). This issue is problematic since teachers' inability to assess their teaching would have detrimental effects on their teaching, students, and, eventually, education. Thus, more attention should be given to the development of teachers' LAL.

Previous studies on ESL and EFL teachers' LAL revealed different results. ESL teachers have been reported to have good LAL. They have relatively sound and clear knowledge of assessment and its implementation (Singh, Singh, et al., 2022). ESL teachers are also reported to be able to use multiple assessment strategies when assessing their students (López, 2023; Singh, Muhammad, et al., 2022); therefore, they are considered more ready to adjust their assessment practices in response to the curriculum reform in their teaching context (Joachim & Hashim, 2021). On the other hand, studies on EFL teachers indicated that many EFL teachers still have problems in conducting assessments due to the inadequate LAL. For example, research by Watmani et al. (2020) reported that many EFL teachers were unfamiliar with LAL principles, making it difficult for them to conduct assessment-related tasks in their teaching. Latif's (2021) research also indicated that in many EFL contexts, EFL teachers' LAL is either developing extremely slowly or failing to fulfill the required curriculum standards. EFL teachers have also been reported to be underprepared when facing language assessment-related tasks despite having participated in pre-service education programs (Jan-Nesar M et al., 2020). Similarly, Vogt and Tsagari (2014) revealed that EFL teachers had low knowledge of LAL, and they suggested that investigations on EFL teachers' LAL should be conducted.

The urgency of investigating EFL teachers' LAL becomes more apparent when considering the current number of EFL teachers globally. In the past two decades, we have been observing the rise of globalization in conjunction with postcolonial trends, resulting in a profound impact on the spread of English and English language teaching (ELT) in the world (Cameron & Galloway, 2019), making English as an international language (EIL). In relation to the profusion of EIL, Tajeddin and Pakzadian (2020) reported that English has been the dominant foreign language in the institutional curricula as well as in foreign language teaching and

learning systems. Such a global status of English has brought about potential markets in various areas such as textbook publishing and teaching professions, particularly in the Outer and Expanding Circles, a model of diffusion of English proposed by Kachru to refer to ESL and EFL contexts respectively (in Tajeddin & Pakzadian, 2020). EIL has then challenged the notions of native speakerism in the developments of ELT textbooks and in the preparations of English teachers since the concept of EIL offers possible different norms to exist for the use of English in different varieties of world Englishes around the globe (Matsuda, 2019). In regard to English teachers in particular, the ever-growing number of English speakers in the Outer and Expanding Circle countries has consequently demanded more and more ESL as well as EFL teachers. The global prevalence of students and teachers, more specifically in the Expanding Circle, thus underscores the extensive reach and critical importance of this demographic. When research on LAL has proven that ESL teachers performed better than EFL teachers, implying that they are comparatively more equipped and supported, collecting information about LAL-related professional development activities among EFL teachers might illuminate the unique challenges EFL teachers face. This concern has inspired this systematic literature review (SLR) on portraying how EFL teachers have been trained in developing their LAL.

Many studies have investigated problems related to EFL teachers' LAL and reported that despite the important role of LAL, EFL teachers claimed the insufficiency of trainings in language assessments for both pre-service teachers (Kremmel & Harding, 2020; Liu & Li, 2020; Ölmezer-Öztürk & Aydın, 2018; Sayyadi, 2022; Sultana, 2019) and in-service teachers (Babaii & Asadnia, 2019; Berry et al., 2019; Bøhn & Tsagari, 2021; Giraldo, 2019; Tian et al., 2021). Some studies showed that research projects had positive effects on the teachers' LAL (Baker & Riches, 2017; Giraldo & Murcia, 2019; Kremmel et al., 2018; Levi & Inbar-Lourie, 2020; Ölmezer-Öztürk, 2021; Saputra et al., 2020; Xie, 2021); however, problems related to EFL teachers' LAL persist. A number of studies reported that training on LAL mainly discussed concepts and theories of LAL without any practical implications for teachers' LAL (Bustamante, 2022; Ölmezer-Öztürk, 2021; Sayyadi, 2022; Sulaiman et al., 2021; Tian et al., 2021). Consequently, such trainings may fail to prepare teachers to play active roles in assessing students (Liu & Li, 2020; Ölmezer-Öztürk & Aydın, 2018; Öz & Attay, 2017; Sayyadi, 2022; Sultana, 2019). Therefore, studies which investigate the training programs targeting EFL teachers' LAL should be conducted.

Research that reported issues in trainings on EFL teachers' LAL mainly required the participants to recall LAL training contents. For example, Sayyadi (2022) investigated in-service teachers' LAL by asking the participants to recall their assessment training when they were pre-service teachers. From the data, the researcher concluded that the assessment training failed to improve the teachers' LAL. A study by Sultana (2019) also reported the same phenomenon where the in-service teachers recalled their assessment training during their undergraduate study to determine the in-service teachers' LAL without discussing the training program model. Thus, whether the training that the teachers reported addressed their LAL was still questionable. Therefore, as most studies recommend, the teacher training or professional development programs should be reformed to achieve the targeted development of EFL teachers' LAL. Based on the recommendations, this present study aimed to review empirical studies on EFL teacher training and professional development programs aiming at developing EFL teachers' LAL. This research employed an SLR method to analyze, synthesize, and summarize previous published studies on training programs to develop EFL teachers' LAL.

The results of this research can provide insights into programs to develop EFL teachers' LAL and can be used to improve programs for developing teachers' LAL. Therefore, this research investigated the models used in the training programs to develop EFL teachers' LAL. The study addresses the following research questions:

1. What is the content of EFL teachers' LAL training?
2. What are the LAL training methods?
3. What are the trends in training to develop EFL teachers' LAL?

Method

This SLR explored the models used in assessment training programs that focus on developing EFL teachers' LAL. This research adopted the PRISMA Group 2015 idea of Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) for reporting our systematic review. The review employed a comprehensive research protocol to ensure transparency and rigor (Jesson et al., 2011; Petticrew & Roberts, 2006). Following the protocols, we 1) defined the research question (RQ), 2) identified keywords and databases, 3) conducted a comprehensive literature search, 4) applied exclusion and inclusion criteria, 5) critically appraised the quality of the sources, and 6) synthesized the studies.

Identification

The researchers identified published articles documents from different databases based on three criteria: i) LAL (mandatory), ii) teacher professional development or teacher training, and iii) articles and proceedings. Research articles written in languages other than English were excluded from this research. This research only covered online databases for broader coverage and easier access to the databases. ERIC, ScienceDirect, Taylor and Francis, SAGE Journals, ProQuest, and Springer databases were explored to find relevant studies. Based on previous studies, LAL seems to have been discussed in the research literature between 2008 and 2009. Hence, the analysis period was set for the last ten years.

Search String

The search string was based on contextualized keywords and contained concepts related to the research question, including synonyms and abbreviations commonly used by researchers in writing article titles. The definition of keywords has been calibrated based on several informal searches to minimize the risk of omission. We restricted the query to titles and keywords to perform the search from the databases. The terms Language Assessment Literacy and Teacher Professional Development are included in the search since they become the focus of this research. Boolean terms "AND" and "OR" (Gough et al., 2017) were used between the keywords in the search string to find sufficient data about Language Assessment Literacy and Teacher Professional Development. Thus, we used the following final search query: ("Language Assessment Literacy" OR "LAL" OR "Assessment Literacy" OR "Language Assessment") AND ("assessment training" OR "Teacher professional development" OR "teacher training").

The generic search string was defined, and wildcard characters were used to cover as many variations of the terms as possible. Filters were also applied in the initial search to narrow down the article selection and focus more on the targeted articles. Filters in the initial search were used for a more straightforward further screening process. Filters applied were for articles written in reviewed journals, written in English, and published between 2012 and 2022. These

filters were applied so that in the initial search, two exclusion criteria, such as articles not written in English and articles published before 2012, were already excluded during the initial search. Thus, the screening process can focus more on topics such as the disciplines covered in the study. The initial search from the six databases using the keywords and the search filters identified a total of 802 articles.

Screening Procedures

The screening procedures were done twice to obtain the final selection of the articles included in the reviews. The first screening was done by the first and second authors on the databases using the inclusion criteria. The first author did the first screening on Taylor & Francis, SAGE Journals, and ProQuest databases. Meanwhile, the second author performed screening procedures on ERIC, Science Direct, and Springer Link databases. The 802 articles were screened by reading the titles and keywords written after the abstract to see if the articles discussed Language Assessment Literacy, LAL, Assessment Literacy, EFL, ELT Assessments, assessment training, and assessment professional development programs. The article abstracts were scanned if the titles and the keywords did not give the expected information. For the initial screening, we documented the results in a Google Spreadsheet. The initial screening using inclusion criteria yielded 176 articles to be included in the second screening.

The first and second authors did the second screening of all 176 articles for final selection. In this process, the first and second authors read the articles' abstracts to determine if the articles discussed EFL/ELT assessment training models or EFL/ELT assessment professional development programs. The first and second authors documented the second screening in a Google Spreadsheet. To finalize the screening process, the first and second authors held a meeting to discuss different selection decisions. The meeting resolved issues on the different selections to finalize the selected articles. Based on the Kappa coefficient analysis, the inter-rater reliability for the article selection was .876, which was categorized as the almost perfect agreement between the raters (Viera & Garrett, 2005), in this case, between the first and the second authors. Of the 176 articles, 28 articles were selected for the review. The review's main focus would be the models used in assessment training or professional development programs to develop EFL teachers' LAL. The article selection and the screening procedures can be seen in Figure 1.

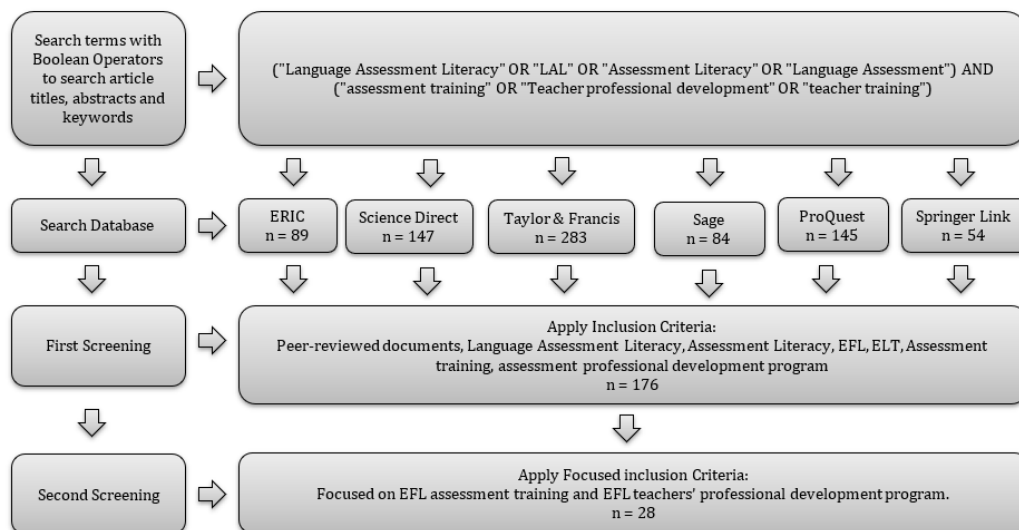


Figure 1. Flowchart of the Selection and Screening Procedure

The Profiles of the Selected Articles

After the final selection, this systematic literature review included 28 articles discussing assessment training models used to develop teachers' LAL. Figure 2 provides information on the number of articles included per year. The figure shows that the earliest articles reviewed in this research were published in 2013, while no article published in 2014 and 2016 was included.

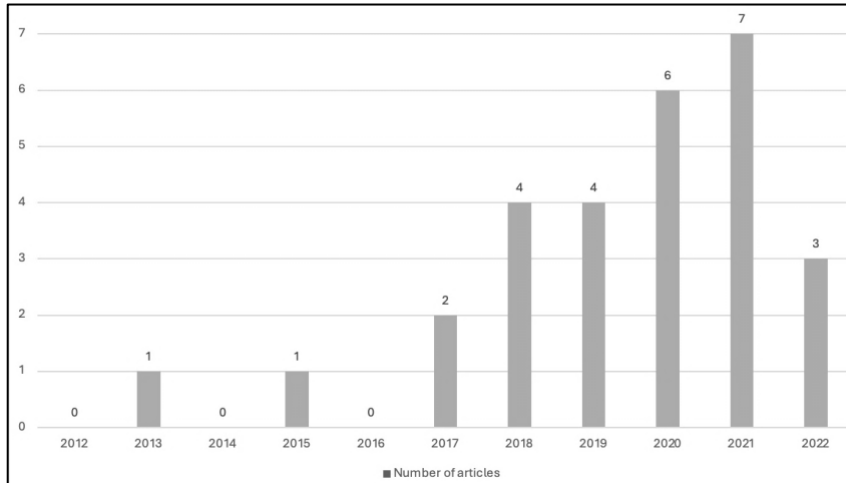


Figure 2. Numbers of Articles Across Publication Years

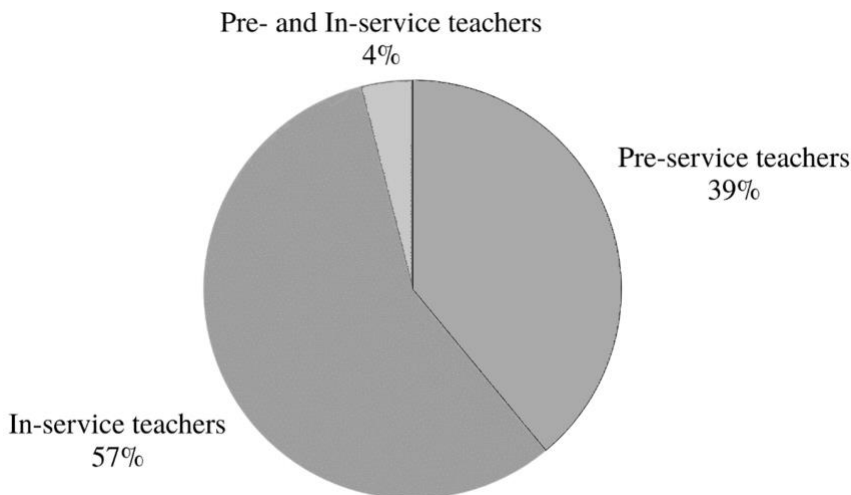


Figure 3. The Contexts of the Research

Figure 3 presents the research contexts where the training programs were implemented. In-service teacher context ranked the highest number (57%), and one study aimed to develop both pre-service and in-service teachers. In addition, Table 1 below provides detailed information on the educational contexts of teachers who participated in the studies.

Table 1. Publication and Context of the Research

In-service	Pre-service	In-service and Pre-service
Babaii and Asadnia (2019)	Giraldo and Murcia (2018)	Giraldo (2021)
Baker and Riches (2017)	Giraldo and Murcia (2019)	
Cinkara (2020)	Ho and Yan (2021)	
Cui et al. (2022)	Kvasova (2022)	
González (2021)	Lam, (2015)	
González et al., (2018)	Lee et al. (2021)	
Janssen (2022)	Levy-Vered and Alhija, (2018)	
Jeong (2013)	Ölmezer-Öztürk, (2021)	
Klug et al. (2018)	Restrepo Bolivar (2020)	
Kremmel et al. (2018)	Torshizi and Bahraman, (2019)	
Lan and Fan (2019)	Xie (2021)	
Levi and Inbar-Lourie (2020)		
Osidak et al. (2021)		
Saputra et al. (2020)		
Shafaghi et al. (2020)		
Yeo (2021)		

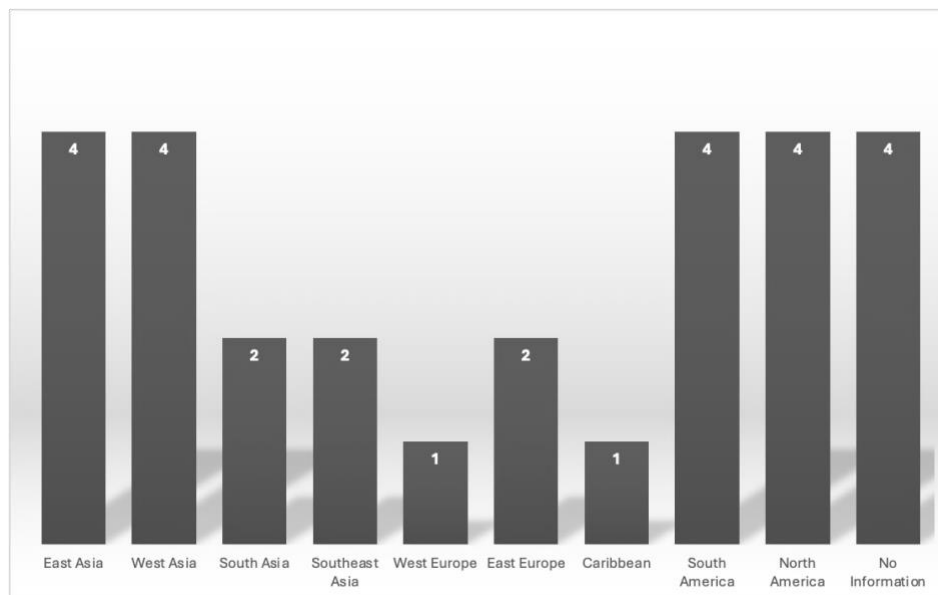


Figure 4. Regional Contexts of the Studies

Figure 4 shows the contexts of the research based on regions where the research participants are from. Eight regions were recorded in the study, with the highest regions being West Asia, South, and North American regions. However, four articles did not state where the research participants are from.

Data Analysis

Thematic analysis was done on the 28 selected articles to find information linked to the research questions. The selected articles were read meticulously and repetitively to see recurrent themes

and patterns, focusing on the training models reported in the article. The reports on the training models were reduced into codes, categories, and themes to generate the review findings. The first author completed coding, categorizing, and thematizing. The results were then verified by the second and third authors to maintain the reliability of the coding.

Findings

Two main themes emerged from the articles after the analysis of the 28 articles. The first theme was the training content that discusses the materials, topics, and issues covered in training. The second theme was the methods used for delivering the training content. In terms of contents, the training models for developing EFL teachers' LAL included an introduction to theories of language assessments, activities involving the practical implications of the language assessment theories, and teachers' involvement in developing assessments. Meanwhile, two main methods for delivering the training content include collaborative and reflective practices. These methods were reported in many studies included in this SLR. The summary of the contents and methods of the training is presented in Figure 5.

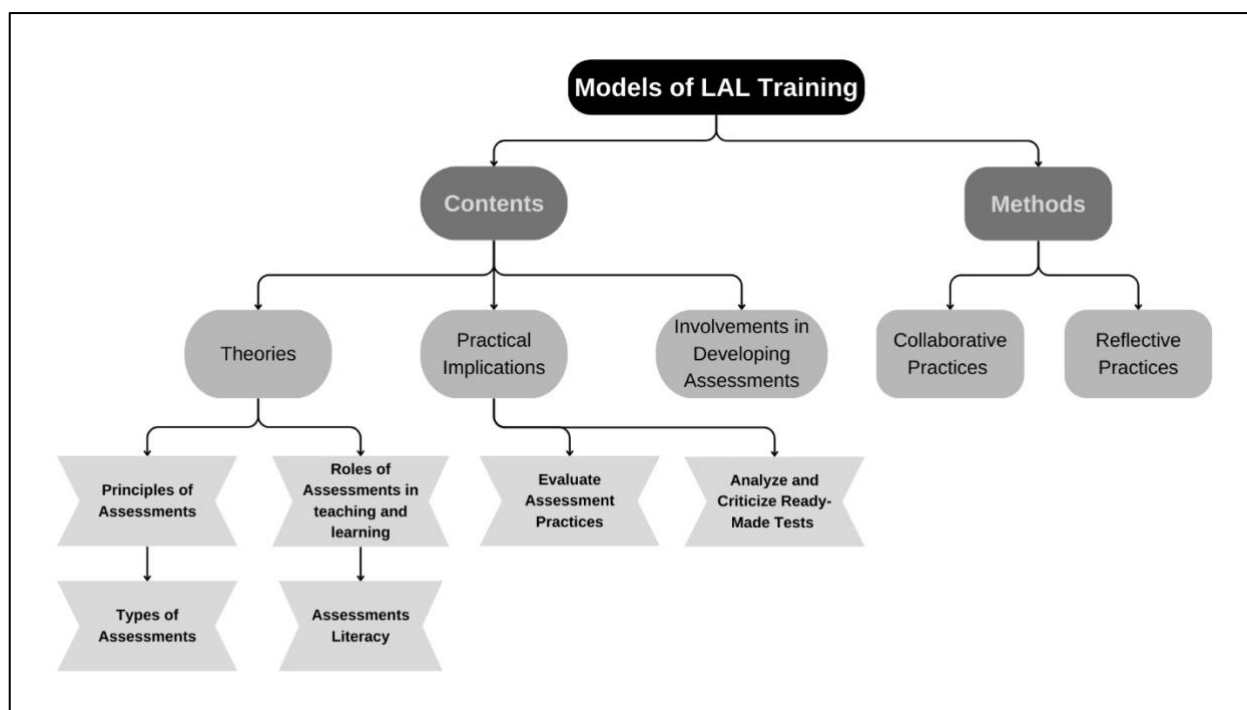


Figure 5. Summary of the findings

The Contents of EFL Teachers' LAL Training

The contents of the training programs presented in the articles include theories of language assessments, practical implications of the language assessment theories, and teachers' involvement in developing assessments. However, the studies put different emphases on those three contents in the training programs. For example, six studies have devoted quite a significant portion to the discussion of language assessment theories, such as studies carried out by Jeong (2013), Lam (2015), Baker and Riches (2017), Levy-Vered and Alhija (2018), Ölmezer-Öztürk (2021), and Kvasova (2022). Most of the trainings were usually done for pre-service teachers to lay a theoretical foundation for language assessment (e.g., Kvasova, 2022; Ölmezer-Öztürk, 2021). Nine studies focused more on the practical implications of the

assessment theories (Babaii & Asadnia, 2019; Giraldo & Murcia, 2019; González, 2021; Osidak et al., 2021; Shafaghi et al., 2020; Yeo, 2021). Ten studies devoted a more extensive discussion on the activities involving teachers in developing language assessments (Cinkara, 2020; Cui et al., 2022; Ho & Yan, 2021; Janssen, 2022; Xie, 2021). However, the other three studies (Klug et al., 2018; Lan & Fan, 2019; Restrepo Bolivar, 2020) did not indicate the proportion of those three elements in the articles. The next part of the findings discusses how each element was incorporated into the training programs will be presented.

The theories of language assessments. The studies reported some theories included in their assessment training. The first theory presented was the principles of language assessments, including tests or assessments validity, reliability, practicality, and authenticity (Baker & Riches, 2017; Cui et al., 2022; Giraldo & Murcia, 2019; Kremmel et al., 2018). The second theory also included in the training was the role of language assessment in language learning. Studies by Levi and Inbar-Lourie (2020) and Ölmezer-Öztürk (2021) included the theory on the role of assessment and evaluation in the learning environment. In addition, studies by Lam (2015) and Ölmezer-Öztürk (2021) included discussions on how assessments can be used to promote students' learning. Thus, the introduction of terms such as Assessment for Learning was also included (Lam, 2015; Levi & Inbar-Lourie, 2020). This theory of the role of assessments may broaden the teachers' knowledge of the uses of assessments beyond monitoring students' achievement. The third theory of assessment discussed the types of assessments, such as large-scale or high-stakes assessments (Cui et al., 2022; Lam, 2015) in contrast to classroom assessments or low-stake assessments (Cui et al., 2022; Jeong, 2013; Lan & Fan, 2019). Other types of assessments that were presented in the theory of assessment were formative and summative assessments (Lam, 2015; Levi & Inbar-Lourie, 2020), and alternative assessments (Babaii & Asadnia, 2019; Jeong, 2013; Ölmezer-Öztürk, 2021). The fourth theory deals with LAL. Interestingly, although the training presented in the articles addressed teachers' LAL, the theory of assessment literacy was not included in the training until 2021, as presented in Ölmezer-Öztürk's (2021) study. Assessment literacy is a new concept, but not many studies included the concept in trainings.

The practical implications of language assessment theories. Besides the theories of language assessments, practical implications of the theories were also used in trainings to engage participants in activities to apply the language assessment theories into practice. The activities were done to strengthen the participants' understanding of the theories. The participants were involved in reflecting on their assessment practices based on the theories taught earlier in training and critiquing and analyzing already available tests or assessments. For example, Babaii and Asadnia (2019), after the participants had learned a language assessment theory from reading research papers, evaluated their current assessment practices based on the theories they had learned from the research papers. The same practice was also reflected in the studies by Saputra et al. (2020) and Yeo (2021). The participants analyzed the validity and reliability of assessment samples (Giraldo & Murcia, 2019). Other studies involved the training participants in analyzing exam questions (Ölmezer-Öztürk, 2021), achievement tests (Cui et al., 2022; Levy-Vered & Alhija, 2018), assessments for measuring writing skills and their rubrics (González, 2021), and tests for national examinations (Baker & Riches, 2017). Both the reflection and analysis were done after the training participants learned theories related to the activity. For example, a study by Giraldo and Murcia (2019) required the participants to analyze the validity and reliability of tests because they included the theory of

validity and reliability of tests. So, the participants had the background knowledge to analyze the tests.

The teachers' involvement in developing assessments. This course content is mostly used in the training programs in developing EFL teachers' LAL. In this activity, the teachers involved in the study were actively involved in developing assessments. Some training required participants to develop assessments as part of the training activities, and the assessments were not administered in any context (González, 2021; Jeong, 2013; Kvasova, 2022; Levy-Vered & Alhija, 2018). Other training programs required the participants to develop assessments that were administered to the participants' teaching context. The assessment can be in the form of tests or performance-based assessments. Some of the training required teachers to design high-stakes tests such as English Proficiency Tests (Ho & Yan, 2021; Kremmel & Harding, 2020), diagnostic tests (Cui et al., 2022), and placement tests (Janssen, 2022). Other training programs involved teachers in projects to develop classroom-based assessments, such as large-scale quizzes for examinations (Cinkara, 2020), classroom tests (Xie, 2021), and authentic classroom assessments (Lan & Fan, 2019; Saputra et al., 2020). As stated above, to be involved in assessment development, the participants underwent a rigorous process starting with planning, development, implementation, and evaluation. In the planning stage, the participants are involved in activities such as writing test specifications and gathering test logistics such as audio for listening tests or reading texts for reading tests (Kremmel et al., 2018) or collecting information such as objectives of the tests (Janssen, 2022), students' English ability and learning progress (Xie, 2021). The development stage involved teachers drafting, reviewing, feedbacking, editing, and revising the assessments (Ho & Yan, 2021; Janssen, 2022; Xie, 2021). Meanwhile, the implementation stage is administering the assessments to the intended audience, such as students at secondary schools (Cinkara, 2020; Kremmel et al., 2018) and students at universities (Cui et al., 2022; Janssen, 2022). The last process reported was evaluation, which involved activities like reflecting on the process of the assessment development or the developed assessments (Kremmel et al., 2018) and receiving constructive feedback from the mentors (Xie, 2021). Although the evaluation process was written as one stage in assessment development, only a few articles discussed how it was done in the research.

The LAL Training Methods

One important theme in the analysis is how the training content was delivered to ensure the effectiveness of the training. The first method that was used in the assessment training was collaborative practices, which was found in nine articles. Collaboration can be done among the training participants (Giraldo & Murcia, 2019; Ho & Yan, 2021; Saputra et al., 2020) with mentor teachers or training facilitators (Cinkara, 2020; Lee et al., 2021; Osidak et al., 2021; Xie, 2021), or with assessment professionals (Baker & Riches, 2017). The collaborative practices also involved the participants in the community of practice, such as the Professional Learning Community (Ho & Yan, 2021; Saputra et al., 2020). The collaboration usually happened during the participants' involvement in developing assessments when they worked together to develop assessments or when the mentor teachers, training facilitators, or assessment specialists gave feedback on the developed assessments during the development or evaluation stage. This collaborative practice is essential in developing teachers' LAL since it was used repeatedly in the articles.

The second method found in the articles reviewed was the use of reflective practices in developing teachers' LAL training. Ten articles mentioned reflection throughout the training

process. The reflections were done in the form of journal and narrative writing (Babaii & Asadnia, 2019; Restrepo Bolivar, 2020), reflective discussion forum (González et al., 2018), and reflection report (Ölmezer-Öztürk, 2021). The reflections also varied in the form of content. For example, participants can write reflections about the current assessment practices (Babaii & Asadnia, 2019; González, 2021; Restrepo Bolivar, 2020; Saputra et al., 2020). In other studies, participants also wrote reflections about the progress they experienced during the training (Kvasova, 2022; Ölmezer-Öztürk, 2021; Restrepo Bolivar, 2020), or reflection on the developed assessments and the implementation of the developed assessments (Klug et al., 2018; Lee et al., 2021; Levi & Inbar-Lourie, 2020). The reflection was also done for participants' future assessment practices and plans after the training program (Kvasova, 2022; Restrepo Bolivar, 2020; Saputra et al., 2020).

The Trends in Training for Developing EFL Teachers' LAL

This research aimed to examine the models for assessment training programs to develop EFL teachers' LAL through an SLR involving 28 previous research articles published between 2012 and 2022. The reviews showed that there had been changes in the trends of LAL training in ten years. The first change was the assessment types that became the focus of training. In the studies published in 2012 – 2017, the training focused on large-scale tests, such as national examinations. For example, a study by Lam (2015) reported a training program that focuses on large-scale tests and theories underpinning the tests and discusses how to make use of the assessment to improve classroom instruction. Other research by Baker and Riches (2017) focused on developing the ability of participants to critique participants' national examination tests and to revise the national examination tests. Meanwhile, studies published in 2019 – 2022 focused more on classroom-based and alternative assessments than tests. For example, Lan and Fan's (2019) research focused on classroom-based LAL, and Levi and Inbar-Lourie's (2020) research focused on developing participants' ability to develop performance-based assessments and the rubrics accompanying the assessments. In addition, more and more training on classroom-based assessment was reported (e.g., Kvasova, 2022; Lee et al., 2021). Meanwhile, the articles published in 2018 focused on large-scale tests (e.g., Kremmel et al., 2018) and classroom-based assessments (e.g., González et al., 2018; Levy-Vered & Alhija, 2018). Publications in 2018 seem to have transitioned from focusing on large-scale tests to focusing more on classroom-based assessments.

Another gradual change that can be seen from the articles reviewed was the nature of the training participants' involvement in developing assessments. In the earlier years of publications, the participants were developing assessments for the training only, and the developed assessments were not used in a real teaching context (e.g., Giraldo & Murcia, 2018; Levy-Vered & Alhija, 2018; Torshizi & Bahraman, 2019). Meanwhile, training reported in later years involved teachers in developing assessments in their teaching context, which also involved more complex processes such as reviewing, revising, administering, and evaluating the developed assessments (Cinkara, 2020; Cui et al., 2022; Janssen, 2022; Xie, 2021). The change may reflect more consideration of the participants' assessment contexts to bridge the gap between the training and the assessment context. Thus, collaboration in developing assessment also changes from collaboration among participants, between participants and training facilitators, to collaboration between participants and mentor teachers who were not training facilitators, and between participants and assessment specialists. The data showed a more complex collaboration among stakeholders.

Discussion

This research revealed models of assessment training programs to develop EFL teachers' assessment literacy. The models comprised the contents of the training and the methods in which the training was conducted. This review revealed that assessment training should cover both theoretical frameworks of assessment, including theories of principles of language assessments that discuss the validity, reliability, authenticity, and practicality of language assessment. The reviewed articles also discussed the theories of the roles of language assessments in teaching, types of assessments, and assessment literacy. These theories of language assessments can enhance teachers' knowledge of the basic principles of language assessments, which become the foundation of teachers' sound assessment practices (Stiggins, 2014). This finding is also in line with Latif and Wasim's (2022) research which revealed that teachers lack knowledge of assessment concepts such as principles of language assessments and different types of assessments. Besides, teachers' knowledge of language assessment theories also influences their assessment practices. Therefore, assessment training should include a theoretical basis for language assessment (Chan & Luk, 2022; Latif & Wasim, 2022). In addition, the inclusion of LAL theories in assessment training also aligns with the suggestion of a study by Watmani et al. (2020) that reported the lack of knowledge of assessment literacy principles by teachers, which may hinder teachers from performing assessment-related tasks.

The following important content of assessment training reported in the articles was the practical implication of assessment theories and the involvement in assessment development. This training content is equally crucial to teachers' LAL because many previous studies reported teachers' inability to use assessment theories in their assessment practices (Gan & Lam, 2022; Sayyadi, 2022; Watmani et al., 2020). Therefore, helping teachers to see the practical implications of the assessment theories is essential for developing teachers' LAL. This research finding highlights similar issues raised by Sulaiman et al. (2021), who argue that both practical and theoretical components of language assessment need to be incorporated into assessment training programs addressing teachers' LAL. Even though Sulaiman et al. (2021) proposed more emphasis on the theories of language assessment than the assessment practices, they agreed that both practical and theoretical exposure to language assessments are essential. Besides, some research recommended activities on helping teachers learn about the practical implications of assessment theories, reflecting on the participants' past assessment practices, which was also proposed by Gan and Lam (2022) to better assist the teachers toward new and reformed assessment practices. Another activity presented in the articles was the participants' involvement in developing assessments, which was also sought as an important activity in assessment training to expand teachers' LAL (Gan & Lam, 2022; Sayyadi, 2022; Stiggins, 2014). Sayyadi (2022) argued that teachers still need training that involves them in the development of assessment. According to Stiggins (2014), equipping teachers with the skill to develop sound assessments reduces the risk of using inappropriate assessment methods that will give false information on students' progress and the teaching and learning process.

Another finding discussed in the reviewed articles was the training methods used to deliver the training content. Across the 28 articles, collaborative and reflective practices were presented as the methods used in training. Gan and Lam (2022) agree that collaboration enforces the interaction between colleagues, trainers, mentor teachers, assessment specialists, and other stakeholders to build a shared understanding of LAL acquisition. Bøhn and Tzagari (2021) and Rohmah (2018) also emphasize that collaborating with multiple stakeholders, such as

colleagues, school management, students, and parents, can create fruitful assessment cultures, which is an important dimension in teachers' LAL. Ahmadi and Ketabi (2020) also proposed that collaboration between teachers and students is an important element of LAL development. While collaboration with students and parents was deemed important (Ahmadi & Ketabi, 2020; Bøhn & Tsagari, 2021), the present study did not find any information about such collaboration. It may be worth noting that they should be included in the assessment training discussion.

Reflective practices were also reported as a method used in assessment training in the reviewed articles. The findings revealed that the training participants had to reflect on the assessment and the training itself. The training participants reflected on their past assessment practices, their developed assessments, and their future assessment practices. They also did reflections on their progress in the training where they could share the essential knowledge and skills they learned from the training and the plan after training related to their assessment and teaching. A study by Tian et al. (2021) reported that through reflective practices, teachers developed more awareness of their assessment practices to design meaningful tasks and develop effective assessments. Gan and Lee (2016) also argued that reflection is an important tool that enables teachers to capture their teaching experiences, including the assessment, and tap into their needs, issues, and concerns in their teaching. This reflection is important to shape their understanding of their practices because when they do reflections, they can express their understanding of their teaching practices, including their assessment practices (Gan & Lee, 2016).

The last finding of this study is the trends in assessment training programs. The findings related to the trends of the type of assessment included in the training were in line with DeLuca et al. (2016), who argue that there is a growing emphasis on the themes of alternative assessments instead of standardized tests. Fitriyah et al. (2022) also suggested paying more attention to classroom-based assessment for teachers' professional development programs. Watmani et al. (2020) also acknowledge trends from large-scale testing to smaller-scale assessment practices. More emphasis on the classroom-based assessment was also suggested by Stiggins (2014), who argues that teachers should be able to develop assessments that can give accurate information about the classroom teaching and learning process, and large-scale tests cannot do that. Meanwhile, the progression of the nature of collaboration has been discussed in Gan and Lam's (2022) study that addresses the complexity of collaboration among stakeholders involved in assessments.

Future Research on Developing EFL Teachers' LAL

The study has some implications for future research on EFL teachers' LAL. First, considering the small number of articles related to training programs for developing teachers' LAL, future research on such training programs is still needed. For pre-service EFL teacher training programs, research on the effectiveness of the existing assessment course to develop pre-service teachers' LAL can be conducted to evaluate the existing training. The result of the evaluation of the assessment can be data to design more effective training programs. For in-service EFL teachers, with the shifting trends from standardized tests to classroom-based assessments, future research can investigate the training programs to equip these teachers for this responsibility better. Second, since the trend of assessment training is in favor of collaboration among parties involved in assessments, such as among trainees, trainers, assessment practitioners, and assessment experts, research on the collaboration models that are

effective for developing EFL teachers' LAL can be carried out to better design assessment training.

Finally, due to the increasing trends of digital or online assessment in some educational contexts, future research can also be devoted to investigating EFL teachers' LAL for these types of assessments. Research can also investigate how to train EFL teachers for these types of assessments. The articles reviewed in the current study have not discussed digital/online assessments. Therefore, it is worth noting that there is a wide range of topics that researchers can explore regarding digital/online assessments.

Conclusion

This SLR investigated language assessment training models to develop EFL teachers' LAL. Using the PRISMA Model 2015, a careful selection process involving identifying online journal databases, determining search strings, and screening procedures was carried out to find articles suitable for review to answer the research questions. The review included and analyzed twenty-eight articles to find themes related to the assessment training models. The findings revealed that assessment training targeting the development of teachers' LAL covered both theoretical and practical aspects of language assessment. Theoretical aspects were included to give a basic understanding of language assessment as the foundation of teachers' assessment practices. Theories such as principles of language assessments, the roles of assessments in education, assessment types, and assessments literacy were introduced in training. The practical implication of the theories was also included in the training to help training participants apply the theories to their assessment practice. Participants were engaged in activities such as analyzing assessments, reflecting on past assessment practices, and developing assessments. The training was done using methods that required participants to have hands-on experience throughout the programs. Collaboration and reflection were also important components of assessment training. These two methods were reported to affect the participants' LAL positively. Finally, there is also a growing trend where assessment training programs moved towards developing teachers' competencies in designing classroom-based assessments.

This research offers crucial insight into the training models to improve teachers' LAL. The training model could be implemented for both pre-service and in-service EFL teachers with a more significant incline toward the in-service teachers. The training model can be adapted to teacher training programs or professional development. Although courses or training on LAL and teachers' professional development programs focusing on LAL may be less than those programs for developing teaching competence, with the use of the appropriate model, the training could be designed to develop EFL teachers' LAL effectively. In addition, the trends of assessment training move toward teachers' classroom assessment practice, indicating the importance of considering assessment context in designing assessment training. The limitation of this research is that the source of data did not include major databases such as Scopus and Web of Science. While the databases used for this research covered most publications in the education field, the use of major databases may yield more articles that can be included in the study. The number of articles included in this SLR can be considered small for SLR, which means that the research on LAL, especially the ones focusing on training models to develop teachers' LAL, still needs to be improved. Therefore, future research on LAL training programs is still widely open. Future research on assessment training and its effects on teachers' LAL can still be explored.

About the Authors

Indah Puspawati is currently a doctorate student at Department of English, Faculty of Letters, Universitas Negeri Malang, Indonesia. She is a faculty member at the English Language Education Department of Universitas Muhammadiyah Yogyakarta, Indonesia. Her research interests include language assessment literacy, democratic language assessment, and student voices in language assessments. ORCID ID: 0000-0002-9577-1320

Maharani Khansa is a doctorate student at Department of English, Faculty of Letters, Universitas Negeri Malang, Indonesia. She is an English teacher for Junior High School students in Batu, East Java. Her research interests include extensive reading, students and teachers' literacy, and reading engagement. ORCID ID: 0000-0002-9714-0012

Utami Widiati is a faculty member at Department of English, Faculty of Letters, Universitas Negeri Malang (UM), Indonesia. She has researched and published extensively in the areas of TEFL, curriculum and material development, SLA, and teacher professional development. She is a pre-service and in-service teacher trainer and has written secondary school English textbooks prescribed by the Indonesian government. Currently, she is the chief editor of TEFLIN Journal and is the TEFLIN President, The Association for the Teaching of English as a Foreign Language in Indonesia. ORCID ID: 0000-0002-8603-4556.

Acknowledgement

The first author would like to express her gratitude to the Center for Higher Education Funding (BPPT) and Indonesian Endowment Fund for Education (LPDP) from the Ministry of Finance Republic Indonesia for granting the scholarship and supporting this research.

To Cite this Article

Puspawati, I., Khansa, M., & Widiati, U. (2024). Developing EFL teachers' language assessment literacy: A systematic literature review on teacher training programs. *English as a Second Language Electronic Journal (TESL-EJ)*, 28(2). <https://doi.org/10.55593/ej.28110a8>

References

- Ahmadi, M. R. S., & Ketabi, S. (2020). Features of language assessment literacy in Iranian English language teachers' perceptions and practices. *Journal of Teaching Language Skills*, 38(1), 191–223. <https://doi.org/10.22099/jtls.2020.34843.2739>
- Ahmadi, S., Ghaffary, S., & Shafaghi, M. (2022). Examining teacher assessment literacy and instructional improvement of Iranian high school teachers on various fields of study. *International Journal of Language Testing*, 12(1). DOI: 10.22034/IJLT.2022.146981
- Aria, D., Sukyadi, D., & Kurniawan, E. (2021). Teacher assessment literacy: Indonesian EFL teachers' self-perceived on classroom-based assessment practice. *English Review: Journal of English Education*, 10(1). <https://doi.org/10.25134/erjee.v10i1.5349>
- Babaii, E., & Asadnia, F. (2019). A long walk to language assessment literacy: EFL teachers' reflection on language assessment research and practice. *Reflective Practice*, 20(6), 745–760. <https://doi.org/10.1080/14623943.2019.1688779>

- Bahtiar, I., & Purnawarman, P. (2020). Investigating English teachers' comprehension in Language Assessment Literacy (LAL). *Advances in Social Science, Education, and Humanities Research*, 508(Icite), 303–310. <https://doi.org/10.2991/assehr.k.201214.253>
- Baker, B. A., & Riches, C. (2017). The development of EFL examinations in Haiti: Collaboration and language assessment literacy development. *Language Testing*, 1–25. <https://doi.org/10.1177/0265532217716732>
- Berry, V., Sheehan, S., & Munro, S. (2019). What does language assessment literacy mean to teachers? *ELT Journal*, 73(2), 113–123. <https://doi.org/10.1093/elt/ccy055>
- Bøhn, H., & Tsagari, D. (2021). Teacher educators' conceptions of language assessment literacy. *Journal of Language Teaching and Research*, 12(2), 222–233. <https://doi.org/10.17507/jltr.1202.02>
- Bustamante, A. (2022). Language assessment literacy of middle school English teachers in Mexico. *Languages*, 7(32), 1–20. <https://doi.org/10.3390/languages7010032>
- Cameron, A., & Galloway, N. (2019). Local thoughts on global ideas: Pre- and in-service TESOL practitioners' attitudes to the pedagogical implications of the globalization of English. *RELC Journal*, 50(1), 149–163. <https://doi.org/10.1177/0033688218822853>
- Chan, C. K. Y., & Luk, L. Y. Y. (2022). A four-dimensional framework for teacher assessment literacy in holistic competencies. *Assessment and Evaluation in Higher Education*, 47(5), 755–769. <https://doi.org/10.1080/02602938.2021.1962806>
- Cinkara, E. (2020). Quiz study as a professional development activity for tertiary-level EFL test writers. *TESL-EJ*, 24(3). <https://tesl-ej.org/wordpress/issues/volume24/ej95/ej95a5/>
- Cui, Y., Liu, Y., Yu, H., & Gao, Y. (2022). Developing English teachers' language assessment literacy in an EAP reform context through test design: A case study. *System*, 109. <https://doi.org/10.1016/j.system.2022.102866>
- Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing*, 25(3), 327–347. <https://doi.org/10.1177/0265532208090156>
- DeLuca, C., LaPointe-McEwan, D., & Luhanga, U. (2016). Teacher assessment literacy: A review of international standards and measures. *Educational Assessment, Evaluation and Accountability*, 28(3), 251–272. <https://doi.org/10.1007/s11092-015-9233-6>
- Firoozi, T., Razavipour, K., & Ahmadi, A. (2019). The language assessment literacy needs of Iranian EFL teachers with a focus on reformed assessment policies. *Language Testing in Asia*, 9(1). <https://doi.org/10.1186/s40468-019-0078-7>
- Fitriyah, I., Masitoh, F., & Widiati, U. (2022). Classroom-based language assessment literacy and professional development need between novice and experienced EFL teachers. *Indonesian Journal of Applied Linguistics*, 12(1), 126–136. <https://doi.org/10.17509/ijal.v12i1.46539>
- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113–132. <https://doi.org/10.1080/15434303.2011.642041>

- Gan, L., & Lam, R. (2022). A review on language assessment literacy: Trends, Foci, and Contributions. *Language Assessment Quarterly*.
<https://doi.org/10.1080/15434303.2022.2128802>
- Gan, Z., & Lee, F. K. J. (2016). Understanding ESL student teachers' learning of classroom practices in the practicum: A case study in Hong Kong. *Asia-Pacific Education Researcher*, 25(2), 251–266. <https://doi.org/10.1007/s40299-015-0258-x>
- Giraldo, F. (2019). Language assessment practices and beliefs: Implications for language assessment literacy. *How*, 26(1), 35–61. <https://doi.org/10.19183/how.26.1.481>
- Giraldo, F. (2020). A post-positivist and interpretive approach to researching teachers' language assessment literacy. *Profile: Issues in Teachers' Professional Development*, 22(1), 189–200. <https://doi.org/10.15446/profile.v22n1.78188>
- Giraldo, F. (2021). Language assessment literacy and teachers' professional development: A review of the literature. *Profile: Issues in Teachers' Professional Development*, 23(2), 265–279. <https://doi.org/10.15446/profile.v23n2.90533>
- Giraldo, F., & Murcia, D. (2018). Language assessment literacy for pre-service teachers: Course expectations from different stakeholders. *Gist Education and Learning Research Journal*, 16, 56–77. <https://doi.org/10.26817/16925777.425>
- Giraldo, F., & Murcia, D. Q. (2019). Language assessment literacy and the professional development of pre-service language teachers. *Colombian Applied Linguistics Journal*, 21(2), 243–259. <https://doi.org/10.14483/22487085.14514>
- González, E. F. (2021). The impact of assessment training on EFL writing classroom assessment: Voices of Mexican university teachers. *Profile: Issues in Teachers' Professional Development*, 23(1), 107–124. <https://doi.org/10.15446/profile.v23n1.85019>
- González, E. F., Alicia, N., & Lopez, V. (2018). Exploring Mexican EFL elementary school teachers' perceptions of online language assessment training. In *Journal of Applied Linguistics and Language Research*. 5(1), 225-2241.
<https://www.jallr.com/index.php/JALLR/article/view/762>
- Gough, D., Oliver, S., & Thomas, J. (2017). *An introduction to systematic reviews* (2nd Edition). Sage Publication.
- Ho, E. C., & Yan, X. (2021). Using community of practice to characterize collaborative essay prompt writing and its role in developing language assessment literacy for pre-service language teachers. *System*, 101. <https://doi.org/10.1016/j.system.2021.102569>
- Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*, 25(3), 385–402.
<https://doi.org/10.1177/0265532208090158>
- Jan-Nesar M, Q., Khodabakhshzadeh, H., & Motallebzadeh, K. (2020). Assessment literacy of Iranian EFL teachers: A review of recent studies. *Journal of Asia TEFL*, 17(2), 689–698.
<https://doi.org/10.18823/asiatefl.2020.17.2.27.689>

- Janssen, G. (2022). Local placement test retrofit and building language assessment literacy with teacher stakeholders: A case study from Colombia. *Language Testing*, 39(3), 377–400. <https://doi.org/10.1177/02655322221076153>
- Jeong, H. (2013). Defining assessment literacy: Is it different for language testers and non-language testers? *Language Testing*, 30(3), 345–362. <https://doi.org/10.1177/0265532213480334>
- Jesson, J., Matheson, L., & Lacey, F. M. (2011). *Doing your literature review: Traditional and systematic techniques*. Sage Publication.
- Joachim, A. & Hashim, H. (2021) ESL teacher's knowledge and readiness on the implementation of School-Based Assessment (SBA) in Malaysian primary school. *Creative Education*, 12, 1066-1078. doi: 10.4236/ce.2021.125079.
- Klug, J., Schultes, M. T., & Spiel, C. (2018). Assessment at school – Teachers' diary-supported implementation of a training program. *Teaching and Teacher Education*, 76, 298–308. <https://doi.org/10.1016/j.tate.2017.10.014>
- Kremmel, B., & Harding, L. (2020). Towards a comprehensive, empirical model of language assessment literacy across stakeholder groups: Developing the language assessment literacy survey. *Language Assessment Quarterly*, 17(1), 100–120. <https://doi.org/10.1080/15434303.2019.1674855>
- Kremmel, B., Eberharter, K., Holzknecht, F., & Konrad, E. (2018). Fostering language assessment literacy through teacher involvement in high-stakes language testing. In D. Xerri & P. Vella Briffa (Eds.), *Teacher involvement in high-Stakes Language Testing*, (pp. 173–194). Springer International Publishing. <https://doi.org/10.1007/978-3-319-77177-9>
- Kvasova, O. (2022). The impact of student teachers' pre-existing conceptions of assessment on the development of language assessment literacy within an LTA course. *Languages*, 7(1). <https://doi.org/10.3390/languages7010062>
- Lam, R. (2015). Language assessment training in Hong Kong: Implications for language assessment literacy. *Language Testing*, 32(2), 169–197. <https://doi.org/10.1177/0265532214554321>
- Lan, C., & Fan, S. (2019). Developing classroom-based language assessment literacy for in-service EFL teachers: The gaps. *Studies in Educational Evaluation*, 61, 112–122. <https://doi.org/10.1016/j.stueduc.2019.03.003>
- Latif, M. W. (2021). Exploring tertiary EFL practitioners' knowledge base component of assessment literacy: Implications for teacher professional development. *Language Testing in Asia*, 11(19), 1–22. <https://doi.org/10.1186/s40468-021-00130-9>
- Latif, M. W., & Wasim, A. (2022). Teacher beliefs, personal theories and conceptions of assessment literacy—a tertiary EFL perspective. *Language Testing in Asia*, 12(11). 022. <https://doi.org/10.1186/s40468-022-00158-5>
- Lee, J., Butler, Y. G., & Peng, X. (2021). Multiple stakeholder interaction to enhance pre-service teachers' language assessment literacy. *Languages*, 6(4). <https://doi.org/10.3390/languages6040213>

- Levi, T., & Inbar-Lourie, O. (2020). Assessment literacy or language assessment Literacy: learning from the teachers. *Language Assessment Quarterly*, 17(2), 168–182. <https://doi.org/10.1080/15434303.2019.1692347>
- Levy-Vered, A., & Alhija, F. N. A. (2018). The power of a basic assessment course in changing pre-service teachers' conceptions of assessment. *Studies in Educational Evaluation*, 59, 84–93. <https://doi.org/10.1016/j.stueduc.2018.04.003>
- Liu, J., & Li, X. (2020). Assessing young English learners: language assessment literacy of Chinese primary school English teachers. *International Journal of TESOL Studies*, 2, 36–49. <https://www.tesolunion.org/journal/details/info/7NzIueOTcy/Assessing-Young-English-Learners:-Language-Assessment-Literacy-of-Chinese-Primary-School-English-Teachers>
- López, A. A. (2023). ESL teachers' perceptions of effective classroom assessment feedback. *Language Teaching Research Quarterly*, 37, 91–107. <https://doi.org/10.32038/ltrq.2023.37.04>
- Matsuda, A. (2019). World Englishes in English language teaching: Kachru's six fallacies and the TEIL paradigm. *World Englishes*, 38(1–2), 144–154. <https://doi.org/10.1111/weng.12368>
- Mellati, M., & Khademi, M. (2018). Exploring teachers' assessment literacy: Impact on learners' writing achievements and implications for teacher development. *Australian Journal of Teacher Education*, 43(6), 1–18. <https://doi.org/10.14221/ajte.2018v43n6.1>
- Ölmezer-Öztürk, E. (2021). Developing language assessment literacy of EFL pre-service teachers through classroom assessment course. *International Journal of Contemporary Educational Research*. <https://doi.org/10.33200/ijcer.932721>
- Ölmezer-Öztürk, E., & Aydın, B. (2018). Toward measuring language teachers' assessment knowledge: Development and validation of language assessment knowledge scale (LAKS). *Language Testing in Asia*, 8(1). <https://doi.org/10.1186/s40468-018-0075-2>
- Osidak, V., Drahinda, O., & Kvasova, O. (2021). Training the trainers in language assessment via mentoring: Building expertise to promote language assessment literacy of Ukrainian university teachers. *Languages*, 6(194). <https://doi.org/10.3390/languages6040194>
- Öz, S., & Attay, D. (2017). Turkish EFL instructors' in-class language assessment literacy: Perceptions and practices. *ELT Research Journal*, 6(1), 25–44. <https://dergipark.org.tr/tr/pub/eltrj/issue/28779/307996>
- Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences: A practical guide*. Blackwell Publishing.
- Razavipour, K., & Rezagah, K. (2018). Language assessment in the new English curriculum in Iran: Managerial, institutional, and professional barriers. *Language Testing in Asia*, 8(1). <https://doi.org/10.1186/s40468-018-0061-8>
- Restrepo Bolivar, E. M. (2020). Monitoring pre-service teachers' language assessment literacy development through journal writing. *Malaysian Journal of ELT Research*, 17(1), 38–52. <https://meltajournals.com/index.php/majer/article/view/558>

- Rohmah, Z. (2018). Enhancing English teachers' professional development: Portraying a mentoring program. *TEFLIN Journal*, 29(1), 90–107. <https://doi.org/10.15639/teflinjournal.v29i1/90-107>
- Saputra, E. R., Hamied, F. A., & Suherdi, D. (2020). The development of beliefs and practices of language assessment literacy: Does a professional learning community help? *Journal of Education for Teaching*, 46(3), 414–416. <https://doi.org/10.1080/02607476.2020.1761250>
- Sayyadi, A. (2022). In-service university-level EFL instructors' language assessment literacy and training needs. *Profile: Issues in Teachers' Professional Development*, 24(1), 77–95. <https://doi.org/10.15446/profile.v24n1.93676>
- Sevimel-Sahin, A., & Subasi, G. (2019). An overview of language assessment literacy research within English language education context. *Journal of Theoretical Educational Science*, 12(4), 1340–1364. <https://acikerisim.aku.edu.tr/xmlui/handle/11630/9903>
- Shafaghi, M., Estaji, M., & Tabataba', A. (2020). Teacher evaluation training and its effect on EFL teacher evaluators' perceptions and practices. *MEXTESOL Journal*, 44(4). https://mextesol.net/journal/index.php?page=journal&id_article=21883
- Singh, C. K. S., Muhammad, M. M., Mostafa, N. A., Yunus, M. M., Noordin, N., & Darm, R. (2022). Exploring ESL teachers' alternative assessment strategies and practices in the classroom. *Journal of Language and Linguistic Studies*, 18(1), 411–426. <https://jlls.org/index.php/jlls/article/view/2867>
- Singh, C. K. S., Singh, H. K. J., Singh, T. S. M., Moneyam, S., Abdullah, N. Y., & Zaini, M. F. (2022). ESL teachers' assessment literacy in classroom: A review of past studies. *Journal of Language and Linguistic Studies*, 18(Special issue 1), 1–17. <https://www.jlls.org/index.php/jlls/article/view/2868>
- Stiggins, R. (2014). Improve assessment literacy outside of schools too. *The Phi Delta Kappan*, 96(2), 67–72. <https://www.jstor.org/stable/24376163>
- Sulaiman, D., Marandi, S. S., & Tajik, L. (2021). The status quo of language assessment literacy among Syrian EFL teachers. *Applied Research on English Language*, 10(2), 33–60. <https://doi.org/10.22108/are.2021.126136.1657>
- Sultana, N. (2019). Language assessment literacy: an uncharted area for the English language teachers in Bangladesh. *Language Testing in Asia*, 9(1), 1–14. <https://doi.org/10.1186/s40468-019-0077-8>
- Tajeddin, Z. & Pakzadian, M. (2020). Representation of inner, outer and expanding circle varieties and cultures in global ELT textbooks. *Asian-Pacific Journal of Second and Foreign Language Education*, 5(10). <https://doi.org/10.1186/s40862-020-00089-9>
- Tian, W., Louw, S., & Khan, M. K. (2021). COVID-19 as a critical incident: Reflection on language assessment literacy and the need for radical changes. *System*, 103. <https://doi.org/10.1016/j.system.2021.102682>
- Torshizi, M. D., & Bahraman, M. (2019). I explain, therefore I learn: Improving students' assessment literacy and deep learning by teaching. *Studies in Educational Evaluation*, 61, 66–73. <https://doi.org/10.1016/j.stueduc.2019.03.002>

- Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: The Kappa statistic. *Research Series*, 37(5), 360–363. <https://pubmed.ncbi.nlm.nih.gov/15883903/>
- Vogt, K., & Tsagari, D. (2014). Assessment literacy of foreign language teachers: Findings of a European study. *Language Assessment Quarterly*, 11(4), 374–402. <https://doi.org/10.1080/15434303.2014.960046>
- Watmani, R., Asadollahfam, H., & Behin, B. (2020). Demystifying language assessment literacy among high school teachers of English as a foreign language in Iran: Implications for teacher education reforms. *International Journal of Language Testing*, 10(2), 129–144. https://www.ijlt.ir/article_118025.html
- Xie, Q. (2021). Developing pre-service English teachers' assessment literacy through design-based research. *Journal of Education for Teaching*, 47(3), 460–465. <https://doi.org/10.1080/02607476.2021.1906155>
- Yeo, M. A. (2021). “Experiencing theory first-hand was delightful and informative”: Bridging the Theory-Practice Gap in Online Language Assessment Training. *Iranian Journal of Language Teaching Research*, 9(3), 93–116. <https://doi.org/10.30466/ijltr.2021.121078>
- Zulaiha, S., Mulyono, H., & Ambarsari, L. (2020). An investigation into EFL teachers' assessment literacy: Indonesian teachers' perceptions and classroom practice. *European Journal of Contemporary Education*, 9(1), 189–201. <https://doi.org/10.13187/ejced.2020.1.189>

Copyright of articles rests with the authors. Please cite TESL-EJ appropriately.