

Investigating the Influence of Experience on Language Assessment Literacy Development: The EFL Teachers' Voices

*¹Khairani Dian Anisa, ¹Sumardi, ¹Endang Setyaningsih

¹Universitas Sebelas Maret, Indonesia

*Correspondence:

khairani.dianisa@student.uns.ac.id

Submission History:

Submitted: October 14, 2023

Revised: October 28, 2023

Accepted: October 30, 2023



This article is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

Abstract

Research in various educational settings reveals that in-service teachers often demonstrate a limited proficiency in language assessment literacy (LAL). Recognizing this, scholars have been investigating factors that might enhance LAL, particularly focusing on the role of experience. Despite this interest, the specific influence of experiential factors on LAL development remains relatively unexplored. Addressing this gap, our study utilized semi-structured retrospective interviews in a narrative inquiry framework to assess how experience impacts the LAL development of two Indonesian EFL teachers. Both teachers, equipped with master's degrees and relevant assessment training, have differing lengths of teaching experience and work in distinct educational institutions. Our findings indicate that foundational LAL knowledge and skills acquired through language assessment courses (LACs) were further enhanced by additional training and practical application in teaching contexts. Both participants, despite their varied teaching tenures, showed similar patterns of LAL development. This progression appeared influenced by their educational backgrounds, positive attitudes towards assessment, and a shared commitment to refining their assessment practices. The study concludes with recommendations for language educators and EFL teachers and areas for future research.

Keywords: EFL teacher; experience; language assessment literacy

INTRODUCTION

Language assessment literacy (LAL) encompasses a teacher's comprehensive understanding of effective language assessment practices and the ability to apply this knowledge in their teaching context (Butler et al., 2021). According to Davies (2008), LAL consists of three key components: knowledge, skills, and principles. Knowledge involves understanding language learning theories, language testing theories, and contextual factors affecting language assessment implementation. Skills entail the ability to plan, construct, administer, evaluate, and report language assessments, including proficiency in statistical

analysis and technology use in assessment practices. Principles focus on appropriate applications, consequences, equity, and ethical considerations in assessment. In addition, [Giraldo \(2018\)](#) expanded on [Davies's \(2008\)](#) model, suggesting that the knowledge component includes awareness of applied linguistics, theory and concepts, and one's own language assessment context.

Moreover, skills encompass instructional abilities, design skills for language assessment, educational measurement, and technological prowess. The principles element involves awareness of and actions towards critical issues in language assessment. Research on LAL has varied in focus. Several studies have developed instruments to measure LAL ([Kremmel & Harding, 2020](#); [Wang et al., 2023](#); [Xu et al., 2023](#)), while others have investigated the impact of professional development on LAL for in-service ([Lam, 2015](#); [Saputra et al., 2020](#)) and pre-service teachers ([Giraldo & Murcia, 2019](#)). Additionally, LAL research has engaged different stakeholders ([Hutami & Putro, 2023](#); [Kim et al., 2020](#); [Kunnan et al., 2021](#); [Vogt et al., 2020](#); [Vogt & Tsagari, 2014](#)). A notable finding by [Gan and Lam \(2022\)](#) is that the investigation of in-service teachers' LAL levels is a predominant theme in LAL studies.

The field of Language Assessment Literacy (LAL) has seen significant evolution in recent years, prompting a diverse range of studies that explore its various dimensions and the factors influencing its development. [Oo et al. \(2022\)](#) highlighted teachers' insufficient language assessment literacy (LAL) across various contexts. In response, scholars like [Malone \(2013\)](#) and [Taylor \(2013\)](#) have shifted their focus towards practical approaches, emphasizing what teachers can realistically achieve in language assessment and advocating a more positive perspective on LAL development. This change in viewpoint is crucial in understanding how experience and other ongoing factors influence LAL. [Jawhar & Subahi \(2020\)](#), [Sun & Zhang \(2022\)](#), and [Zhang et al. \(2021\)](#) explored these mediating factors in LAL development. Their research suggests that a teacher's educational background, professional training, and practical classroom experience, including language assessment practices, play a significant role in shaping their teaching practices, as posited by [Borg \(2003\)](#). These elements collectively influence a teacher's cognition, which is then reflected in their teaching.

Numerous studies have examined the impact of the experiential dimension on LAL among in-service teachers. [Yan et al. \(2018\)](#) found a substantial influence of experience on the LAL of in-service junior high school teachers, mediating the contextual factors. Similarly, [Crusan et al. \(2016\)](#) investigated ESL instructors' linguistic backgrounds and teaching experience in a writing assessment course for tertiary students. Their findings indicated a significant impact of these factors on instructors' beliefs, practices, and LAL. [Sun & Zhang \(2022\)](#) further examined Chinese EFL lecturers' LAL levels, training needs, and mediating factors in LAL development, uncovering a lack of involvement in language assessment procedures, unfavorable attitudes towards LAL, and a pressing need for LAL training implementation. This research also probed the relationship between experience and the EFL instructors' LAL proficiency.

Furthermore, the current research focuses on understanding how experiential factors impact language assessment literacy (LAL) in teachers, specifically those teaching English as a Foreign Language (EFL) in Indonesia. Previous studies have mainly explored whether the experiences of in-service teachers affect their LAL but have not deeply

examined the degree of this influence. These studies often used [Taylor's \(2013\)](#) framework, which outlines eight dimensions of LAL, including knowledge of theory, technical skills, and pedagogical principles.

To address existing knowledge gaps, our study investigates the extent to which experiences shape Indonesian EFL teachers' LAL development, using [Davies' \(2008\)](#) conceptualization of LAL. This framework is widely accepted and comprises three key elements: knowledge, skills, and principles, frequently referenced in scholarly discussions on LAL. Our research methodology includes examining the schooling, training, and teaching experiences of the participants, following [Borg's \(2003\)](#) approach, to comprehensively understand how these experiences contribute to LAL development. The primary research question guiding this study is: How significantly do experiential factors influence the development of LAL in Indonesian EFL teachers?

METHOD

This research employed narrative inquiry, a methodology aligned with the interpretivist paradigm, which emphasizes exploring individual experiences through their personal stories. This approach is suitable for delving into the nuanced experiences of educators. Moreover, the research involved two Indonesian EFL teachers, Sasa and Fita (pseudonyms chosen by the participants themselves), from different academic institutions. These teachers were selected based on specific criteria relevant to the study's focus: they both had attended language assessment courses and training related to assessment and had a minimum of three years of teaching experience. Sasa and Fita both hold master's degrees. Sasa has participated in training to design test items for national examinations and to create Higher Order Thinking Skills (HOTS)-based test items, while Fita attended training on the current curriculum implementation, including assessment, in 2023. Sasa has 17 years of teaching experience and teaches at a public senior high school in Wonogiri. Fita, with four years of experience, teaches at a public junior high school in Kebumen. Both volunteered for this study, which seeks to understand how their diverse backgrounds and experiences contribute to their LAL development.

To understand the impact of experiential factors on the language assessment literacy (LAL) development of two Indonesian EFL teachers, this study employed semi-structured retrospective interviews. These interviews, conducted in Indonesian for clear communication, were based on [Davies' \(2008\)](#) LAL dimensions, as elaborated by [Giraldo \(2018\)](#). This method was chosen because semi-structured interviews are widely used in language teaching research and allow for in-depth exploration of participants' experiences. The interview protocol included open-ended questions to gather detailed information and explore emerging themes. All interviews were recorded and transcribed verbatim for accurate analysis. Given that narrative inquiry falls under the interpretive paradigm, results are expected to vary due to the emphasis on personal experiences, context dependence, and the potential for multiple interpretations of reality. To establish trustworthiness, this study employed strategies like member checking for credibility, debriefing techniques, and source triangulation for dependability and confirmability. Additionally, a detailed explanation of the study's context was provided to enhance its transferability, aligning with [Lincoln and Guba's \(1985\)](#) guidelines as discussed in [Kivunja and Kuyini \(2017\)](#).

Finally, the study utilized Miles et al. (2014) interactive data analysis to analyze the data. This model consists of four processes: "(1) data collection, (2) data condensation, (3) data display, and (4) conclusion drawing/verifying". The data were collected through semi-structured retrospective interviews. The interviews were then transcribed verbatim to ease the data analysis. Data condensation occurred subsequent to the data collection stage. During this phase, the researchers performed various procedures, such as editing, selecting, and summarizing the data gathered in the initial phase. Subsequently, the outcomes were categorized into themes. During the data display phase, the data were organized into distinct tables and descriptions to facilitate the process of drawing conclusions from the data. In the last stage, the process of developing and confirming a conclusion involved synthesizing and interpreting the acquired and analyzed data.

FINDING AND DISCUSSION

The Influence of Teachers' Academic Experiences on the LAL Development Knowledge

Sasa and Fita acquired fundamental knowledge of LAL through their participation in LAC during their undergraduate and graduate studies. Their comprehension of approaches and techniques to administer language assessments, theories and concepts of language assessments, and basic statistical calculations for analyzing language assessment results were cultivated from their academic experiences. Fita acquired knowledge in various approaches and techniques for administering language assessments during a language assessment course (henceforth LAC). She expressed that teachers might evaluate students' speaking proficiency by employing either storytelling or dialogue, taking into account the approach and technique that best corresponded to the learning objective. In the same vein, Sasa also learned the interrelation of applied linguistics and language assessment, albeit succinctly.

Yes. We also learned about it. I think, depending on the learning objective, we can assess speaking through storytelling or dialogue. (I/F/K/AL/02/04/13)

Yes. I am not sure we were taught it in detail, but it was explained. (I/S/K/AL/02/01/11)

Besides, they had an opportunity to construct their knowledge regarding the theories and concepts of language assessment. Both participants articulated that they learned a lot of theories and concepts of language assessment when they were students through attending an LAC. They learned about feedback as one of the formative assessment forms, for instance. Nonetheless, they articulated that they learned this material at a glance.

Giving feedback is also a sort of evaluation. However, it is not stated further or discussed in length in the course. (I/S/K/TC/02/01/23)

Maybe it was said that feedback is a type of assessment, but I can't remember. (I/F/K/TC/02/04/20)

From this course, they also learned how to calculate basic statistical analyses in relation to language assessment, for example, mean, mode, and median.

Yes. That was the first thing I learned. If I'm not mistaken, we calculated it first using an application. [The participant confirmed that she learned how to compute the mean, median, and mode in college] (I/S/K/TC/02/01/36)

I did that calculation during my undergraduate study. (I/F/K/TC/02/04/36)

Sasa and Fita acquired the knowledge through the LAC to appraise the relevance of language assessment in relation to the context. Considering the relevance of the language assessment to context could enhance the meaningfulness of language assessment for students.

I think I learned from the language assessment course that we had to take into account students' characteristics, such as their learning ability. We should adjust to students' abilities. But I think I only learned about it briefly. (I/S/K/LC/02/01/21)

I recall it was explained to me, but I forgot the details. I recall practicing and creating questions based on the curriculum at that time. (I/F/K/LC/02/04/21)

Skills

According to the interviews, the participants expressed that by attending the language assessment course, they started to develop their skills in aligning curriculum objectives, instructions, and language assessments. Moreover, by completing the assignments, they built their skills in constructing blueprints, designing multiple-choice items, analyzing the test items, and interpreting the test results. From their experiences when they were students, Sasa gained the ability to align a test item and the curriculum applied. Similarly, Fita expressed that she learned to determine the complexity level of the test item by following Bloom's Taxonomy once she had comprehended the learning objective in designing a test item. Thus, they could determine the appropriate instructions for their language assessments based on the curriculum objectives and the intricacy level.

We learned how to develop a test item by practicing; it was a course assignment. We had to design the test item in accordance with the curriculum. We chose one material and then created a test item out of it. (I/S/S/IS/02/01/22)

First, we observed the goal. The level of the test item was then determined using the Bloom Taxonomy. Following that, we created the test item. (I/F/S/IS/02/04/24)

Sasa and Fita attempted to acquire skills in blueprint construction for designing an item test during their undergraduate and graduate studies. Both participants learned the skill to create a test item by completing an assignment in the language assessment course. Prior to creating a test item, they were instructed to develop a blueprint. Constructing a test item blueprint enabled teachers to ensure that the test item aligned with the planned learning outcomes.

We were asked to construct a blueprint for the multiple-choice item. (I/S/S/DS/02/01/29)

We made a blueprint based on the material we had previously chosen. Then, after constructing the blueprint, we design a test item. (I/F/S/DS/02/04/28)

Furthermore, Sasa articulated that after learning to construct a blueprint, she learned how to design a multiple-choice item. She elaborated that the multiple-choice test was given to the students in accordance with the curriculum objective of the blueprint she had created. Meanwhile, Fita gained skills in designing a multiple-choice item from the language assessment courses she attended in undergraduate and graduate programs. Neither Sasa nor Fita had learned how to design other types of assessment, such as essays or short answers, and only learned how to design a multiple-choice item from the language assessment course they attended. They recalled that they needed to take into account several things to design a good multiple-choice item. For instance, the options should be homogeneous and approximately equal in length. Sasa stated that while scoring a multiple-choice item, she adopted a scoring system where a correct response was awarded 1 point, while an incorrect answer received 0 points.

Following the construction of the blueprint, we were asked to design a multiple-choice item and distribute it to students. (I/S/S/DS/02/01/29)

The options should not be too dissimilar to one another; for example, one option should not be lengthier than the others. Then, the types of the options are similar, for example, abcd, and they are all adjectives or nouns. (I/S/S/DS/02/01/32)

In terms of designing assessments, I only learned how to design a multiple-choice. I learned that when determining the score for multiple-choice questions, I could give 1 point for a correct answer and 0 points for an incorrect answer. But I also learned the scoring rubric for each language skill in Brown's book. (I/S/S/DS/02/01/35)

In undergraduate and graduate programs, I learned how to design a multiple-choice item. I haven't learned how to design short-answer questions or essays. (I/F/S/DS/02/04/33)

... for example, how to make a distractor, how to make a question. For example, it is necessary to ensure that the questions are comprehensible, that the options are of equal lengths, and that they belong to the same category, such as nouns. That's what I can recall. (I/F/S/DS/02/04/28)

From the interviews, Sasa and Fita narrated that they acquired the skills to perform educational measurements for language assessment. They gained the skills to compute reliability and validity by applying the knowledge gained from the LAC. Furthermore, they were able to analyze data pertaining to test design, including item difficulty. By acquiring these skills, they would be able to discern the adequacy of the test design and pinpoint the items that required enhancement.

Yes. I recall we were asked to analyze our multiple-choice item after the students completed the test. We needed to analyze the level of difficulty, among other things. (I/S/S/MS/02/01/34)

Yes, as I recall, there were a lot of measurements. As I recall, the application was also to determine whether or not it was valid. I think we needed to analyze each item. (I/S/S/MS/02/01/37)

Yes. I learned how to assess validity and reliability. I recall doing the analysis for this assignment. (I/F/S/MS/02/04/44)

Both Sasa and Fita learned to utilize technology in their language assessment practices as a result of taking the language assessment course. They recalled acquiring the skill of utilizing an application to analyze test items. In addition, Fita mentioned that the lecturer provided instructions on how to use the application prior to instructing the students to analyze the test items. Sasa also learned to analyze the data of the test item she designed using an application. The application could generate analysis, which aided her in interpreting the test item result.

I can't remember the name of the application. All I recall is that I had to enter the data, and the results of the analysis from the data I entered would appear. (I/S/S/TS/02/01/42)

The lecturer taught how to use ITEMAN to analyze the assessment. (I/F/S/TS/02/04/48)

Principles

Sasa and Fita expressed that they had not had the opportunity to learn about the ways to administer language assessment appropriately from their experience attending LAC.

The findings from the interview discovered that the knowledge and skills acquired from an LAC served as a guideline for pre-service teachers to conduct language assessments once they became in-service teachers. This result coincides with [Richards and Lockhart \(2007\)](#), who noted that the experiences of prospective teachers influence their future teaching practices. They built their awareness of theories and concepts of language assessment as well as their skills in planning, constructing, and evaluating test items. The assistance from the teacher and the relevant assignments contributed to their LAL development ([Giraldo & Murcia, 2019](#)). However, from the interviews, the LAC emphasized the knowledge and skills aspect of LAL. [Şahin \(2019\)](#) reported a similar result that most LACs in Turkey emphasized knowledge and skills of language assessment. However, several theories and concepts of language assessment were delivered briefly, such as the material about feedback as a form of assessment.

Moreover, the participants narrated that they only had the opportunity to practice designing a multiple-choice item. The result echoed with [Vogt and Tsagari \(2014\)](#), who discovered that the undergraduate students received the least training on scoring and practicing alternative assessments. Indeed, it is crucial for prospective teachers to receive guidance in certain aspects since they need direct instruction, such as in delivering test results, administering testing, and developing assessment tools ([DeLuca & Klinger, 2010](#)). Furthermore, the interviews revealed that the participants had not acquired the principles aspect of LAL from the LAC they had attended. The consideration that in-service teachers put less emphasis on fairness and ethics in administering language assessment ([Taylor, 2013](#)) might be the rationale for this result. The insufficient material covered in LAC might be rooted in the limited time allocation ([Şahin, 2019](#)).

The Influence of Teachers' Training Experiences on the LAL Development Knowledge

The interview disclosed that Fita gained knowledge regarding language acquisition and its influence on language assessment through training. Fita discussed the application of diagnostic tests to identify students' characteristics prior to teaching in the curriculum. She explained that the type of language assessment practiced in her class should align with her students' characteristics. Nevertheless, since the training focused on designing test items, Sasa did not recall learning how to align applied linguistics and language assessment.

We learn about recounting in English. According to the preliminary diagnostic data, for example, 30% of the students in the class are kinesthetic, followed by 30% who prefer to talk and 40% who prefer to write. We may deduce their interests, talents, and learning styles and utilize this knowledge to choose the type of test for recount material they will take. (I/F/K/AL/02/05/03)

Based on their training experiences, Sasa mostly acquired knowledge about the characteristics of a well-constructed multiple-choice since she participated in a training focusing on designing test items.

I learned about the qualities of a good test and other such things. (I/S/K/TC/02/02/06)

Additionally, Fita gained a deeper understanding of how to distinguish language assessments based on their intended objectives.

We learned about assessment, types of assessment, formative assessment, and summative assessment. We learned that formative assessment was an assessment that aimed to improve the learning process rather than only obtaining results or scores. So, it is done at the start and throughout the learning process. Then, at the end of one or more materials, a summative assessment is performed. It is done to assess students' understanding of the subject they have learned. (I/F/K/TC/02/05/01)

Sasa and Fita gained novel comprehension of the guidelines and policies related to the language assessment context from the trainings they had attended.

For high school students, the typical reading text length was 150 words. That is the national examination standard. (I/S/K/LC/02/02/12)

In fact, not only is the midterm exam optional, but so is the final exam. Actually, everything is left up to the subject teacher. (I/F/K/LC/02/06/39)

Skills

They also learned to align the learning objective, instruction, and assessment from their training experiences. Sasa developed her skill to align the test item she designed with the criteria of an HOTS-based test item through her participation in a training focusing on designing an HOTS-based test item. Besides, Fita learned that the language assessment form and instruction could differ based on the students' characteristics when she became a participant in the training about the current curriculum. Regardless of the many formats of language assessments used, the primary goal of the language assessment was to accurately assess the student's language proficiency in alignment with the curriculum's objectives.

We focused on designing the questions we made at the HOTS level. (I/S/S/IS/02/02/28)

On recount text material, for students who prefer spoken text, for instance, we could assess their spoken recount text ability. We can assess their ability by asking them to write a recount text if the students prefer written recount text. Even though the instructions differ owing to the students' characteristics, as long as the objective is met, it is acceptable for the current curriculum. (I/F/S/IS/02/05/03)

Sasa developed the skill to design multiple-choice items for national examinations and HOTS-based items. She recalled acquiring the skill of constructing a stem for an item and the importance of ensuring equal length for the options while designing a multiple-choice item.

I was also taught to create a stem. (I/S/S/DS/02/02/12)

It was typically stated that the options should be equal, and so forth. (I/S/S/DS/02/02/50)

Nevertheless, neither participant had the opportunity to develop their skills in educational measurement for language assessment from the training they attended. Sasa did not analyze her test items since she participated in training that specifically focused on designing test items for national examinations and HOTS-based test items. Meanwhile, Fita attended a training that addressed the implementation of the current curriculum.

Sasa mentioned that she acquired the skill of utilizing online resources to design her test items during her training. She utilized reliable and authentic sources for creating the test items.

At the time, I looked for it myself, but it had to come from a reliable source, even if it was from the internet. For example, it is not permitted to use Wikipedia. As a result, I looked for texts from respectable and reputable websites. (I/S/S/TS/02/02/15)

Principles

Fita learned how to implement fair language assessment administration by taking into account students' comprehension levels through training. By considering students' comprehension levels, teachers could fairly evaluate their abilities.

For instance, all students received the same content in the Simple Present Tense, but certain students might benefit from receiving Simple Present Tense content with simpler vocabulary. Later on, when they had a deeper understanding of the subject matter, Simple Present Tense material might incorporate more complex vocabulary. (I/F/P/CI/02/05/19)

The results of the current study revealed that the training helped teachers enhance their LAL based on the training topic. Sasa exhibited enhanced LAL in the field of designing test items. The training also equipped her with supplementary knowledge about the area. [Restrepo-Bolivar \(2020\)](#) affirmed that the training programs involving constructing assessment tools contributed to teachers' enhanced LAL. Fita's understanding of the alignment of curriculum and language assessment improved when she participated in curriculum training. She gained a deeper understanding of providing language assessments

in accordance with the present curriculum. This result aligned with Cui et al. (2022), who showed that training helps teachers better their LAL. This improvement arose as Sasa and Fita engaged in the practical implementation of the training material. Moreover, the pieces of training were organized in relation to the context (Al-Waaili, 2018), which promoted better comprehension of the training materials to Sasa and Fita.

The Influence of Teachers' Teaching Experiences on the LAL Development Knowledge

Sasa and Fita chose the language assessment practices based on the most appropriate technique that aligned with the learning objectives from their teaching experiences. Moreover, Sasa noted that in assessing her students, she adopted the elements in the scoring rubric she learned in an LAC. This result indicated that the knowledge they acquired from their academic journey equipped them to administer language assessments.

If the topic lends itself to dialogue, such as transactional, we will engage in it. If the theme is about experience, for example, we encourage them to perform a recount monologue. We conform to the theme. (I/F/K/AL/02/06/16)

Previously, the purpose was for students to develop oral, analytical writing; subsequently, I may use debate or just explain in front of the class. Then, for speaking evaluation, what aspects are evaluated, and so on, I refer to what we were taught in college (I/S/K/AL/02/03/20)

Drawing from their teaching experiences, both participants had a comprehensive understanding of the procedures involved in designing tests and the rubrics for the test.

To begin, carefully define the learning objectives. Then, decide which aspect that will be assessed. For example, in descriptive text, I will examine certain language skills through this material. (I/S/K/TC/02/03/05)

For example, we were taught that in determining the rubric for speaking, we included fluency, pronunciation, and so on. If I want to assess my students' speaking ability, the components will be adjusted and not necessarily follow the theory. (I/F/K/TC/02/06/17)

Regarding teaching practice, Sasa promoted the relevance of a language assessment to her classes. Sasa promoted the need to incorporate the relevant language assessment within her teaching practice.

As a sort of language assessment, the school sister created short films. I requested a copy of the films from the teacher. I later showed the films to my students and asked if they wanted to make similar films. They were motivated to create short films, too. They created some short films, and I invited them to provide feedback to one another. (I/S/K/LC/02/03/17)

Skills

Although the language assessment course provided a brief explanation of feedback, participants offered feedback on student's work in their language assessment practices.

I also provide feedback sometimes. (I/S/S/IS/02/03/13).

The feedback is usually about their strengths and what they need to improve. (I/S/S/IS/02/03/21)

Yes, I always provide feedback, both written and spoken. (I/F/S/IS/02/06/22)

Sasa and Fita stated that during their teaching careers, they offered feedback on their students' performances. Sasa communicated to her students both their strengths and areas for improvement. Fita also provided her students with written or spoken feedback. Through the provision of feedback, they facilitated the improvement of their students' skills. They prepared comprehensive blueprints for summative assessments and tended to design simple blueprints for formative assessments.

I usually make a simple blueprint. It is only like if I want to deliver a material, then I will use this way to assess my students for it. (I/S/S/DS/02/03/06)

I make detailed blueprints for mid-term and final examinations. In making the blueprint, I usually comprehend the learning objective and indicators, and then the number of questions will be determined by the indicators I have set. However, if it is a test for my class, I just create a simple blueprint. For example, if I intend to teach this content, I will assess my students in this manner. (I/F/S/DS/02/06/21)

Sasa and Fita created simple blueprints for their language assessment practices. The concise blueprint encompassed the content and the methods used to assess the proficiency of their students. Fita stated that she created a comprehensive blueprint for the mid-term and final exams. She explained that she comprehended the learning objectives and indications involved in creating a summative assessment blueprint. Subsequently, she assigned the number for the questions based on the indicators.

According to the interviews about their teaching experiences, Sasa and Fita revealed that they were able to infer the strengths and shortcomings of their students by analyzing the language assessment results.

If it's a formative test, however, I do it myself. I mean, I analyze it myself, but only briefly. For example, many students answer this question incorrectly. (I/S/S/MS/02/03/24)

I've never thoroughly examined the assessment, such as assessing the item difficulty and so forth. I usually just ask my students which questions they find hardest to answer, and we discuss it afterward. (I/F/S/MS/02/06/35)

Both participants stated that they recognized the specific question that their students struggled to answer, indicating a lack of comprehension of this item among the students. Sasa identified the challenges her students faced by assessing which item had the highest frequency of incorrect responses. Meanwhile, Fita asked her students to identify which questions they considered challenging and subsequently engaged in a discussion with them to address the answers. Furthermore, the interviews revealed that the participants also share similar experiences in employing technologies for their language assessment practices.

I will try to look for texts from reliable resources. For example, from browsing the internet and finding texts from English newspapers. (I/S/S/TS/02/03/22)

I just look for good and authentic texts or worksheets for my teaching on Google. Or, I open the BBC application if I want more authentic and high-quality texts. However, I chose texts written in Indonesian for students in grades 7 and 8 to ease their understanding of the texts. (I/F/S/TS/02/06/28)

Sasa and Fita utilized the internet to search for relevant resources. Nevertheless, they exclusively chose authentic, relevant, and trustworthy sources for their language assessment practices. Sasa used authentic resources, such as English newspapers, in her language assessment practices. Similarly, Fita incorporated authentic material from an application. Nonetheless, to ease her students' comprehension of the text, she selected the texts written in Indonesian.

Principles

From their teaching experiences, Sasa and Fita attempted to encourage fair language assessment practices by minimizing the possibility of cheating in their classes despite not thoroughly learning about the language assessment principles.

If it's a regular test, I can monitor it. For this reason, I like to print the questions so that I can monitor the test. I prefer this way because we don't know if the students look for the answers online. I distributed two types of questions, A and B, to students. For example, this row gets A, and this row gets B, so I can tell who gets A and who gets B questions. It will be easier to monitor for me. (I/S/P/CI/02/03/32)

I usually do speak or performance-based assessments to avoid cheating for my students. If it is done in writing, they are more likely to cheat. (I/F/P/CI/02/06/31)

Sasa and Fita had different ways of ensuring fair language assessment practices, even though they shared a similar aim to avoid cheating during their language assessment practices. Sasa chose to offer different printed test items. At the same time, Fita favored assessing her students using performance-based assessment. This diversity in conducting language assessment arose from the teacher's preference and capability to assess their students effectively.

Furthermore, Sasa recalled inviting her students to express their thoughts about their approaching language assessment. Sasa facilitated democratic language assessment practice by providing an opportunity for students to express their voices on the language assessment administration. Hence, the students and teacher might reach a consensus and prevent any misinterpretation regarding the language assessment practice.

I played the video showcasing the assignment completed by the sister school, and I asked them if they were able to do the assignment they had watched. (I/S/P/CI/02/03/17)

Teachers enhanced their LAL by engaging in the process of trial and error during the administration of language assessments throughout their teaching careers. From this process, they identified which language assessment practices worked well (Richards & Lockhart, 2007). They utilized and modified previous knowledge acquired through

education and training to suit their own circumstances. Furthermore, teachers attempted to acquire and integrate their existing knowledge of assessment in order to create language assessment practices that were most suitable for their classrooms based on a process of trial and error throughout their language assessment administration. Over the course of their teaching careers, both participants acquired expertise and instincts in conducting language assessments. This progression arose from their repeated practice of language assessment and subsequent contemplation of which practices were suitable for their classes (Richards & Lockhart, 2007; Yan et al., 2018). Moreover, despite aiming for similar goals in their language assessment practices, Sasa and Fita performed in different ways to ensure fair language assessment practices in their class. This variation occurred since both participants had their own preferences in conducting their language assessments (Richards & Lockhart, 2007). These findings are consistent with Borg's (2003) assertion that the experiences gained from education, professional training, and teaching experiences are interconnected, along with the contextual elements that shape a teacher's cognition and influence their teaching practices, including assessment.

The present study also found that the number of years of teaching experience had an insignificant impact on the participants' LAL development. Sasa and Fita have a significant disparity in their years of teaching experience. Nevertheless, they both demonstrated the same level of LAL development. This finding aligns with the research conducted by (Jawhar Subahi, 2020; Sun & Zhang, 2022), which both concluded that the number of years of teaching experience had no significant impact on educators. Nonetheless, this study presents contrasting findings compared to (Crusan et al., 2016) that reported a significant impact of teaching experience on a teacher's LAL. Multiple factors are justified to contribute to the different findings.

Based on the interviews, both participants shared similar experiences when attending a LAC. This course enabled them to comprehend the importance of language assessment, hence facilitating the enhancement of their LAL (Sun & Zhang, 2022). Moreover, they showed positive attitudes toward assessment and were eager to improve their language assessment administration, resulting in them performing similar LAL (Zhang et al., 2021). However, it is important to note that both participants in this study hold a magister's degree and involve two participants. The results may differ if larger participants with various educational backgrounds are included in a study. Besides discovering the teacher's experience and the extent of their experiences affecting their LAL development, the present study revealed that a community of practice, such as an English teacher's forum, served as a medium for teachers to exchange novel ideas and knowledge pertaining to their instructional practices echoing the study from (Saputra et al., 2020).

CONCLUSION

This study investigated the impact of the experiential component on LAL development of two Indonesian EFL teachers using the narrative inquiry method. The participants built their knowledge and skills in language assessment through their experiences attending the language assessment course. Moreover, Sasa further improved her knowledge and skills related to designing language test items, and Fita developed awareness of her own context of language assessment practices by attending training. Both participants developed their LAL similarly despite having a wide gap of years of teaching

experience. The similarity in their educational degree, good attitudes towards assessments, and motivation to enhance their assessment practices could potentially explain this outcome.

The present study indicated that the participants established their foundation of the theories and skills in relation to language assessment through a language assessment course. Therefore, language assessment educators should provide comprehensive theories and concepts of language assessment, including the material about feedback and alternative assessments. Moreover, language assessment educators should provide an opportunity for students to practice assessing different types of assessments, e.g., essays, completions, true-false, etc., and provide scoring rubrics to assess different types of assessments. Additionally, the material about LAL should be discussed in an LAC to raise students' awareness of LAL, particularly the way to administer language assessment (principles). Language assessment educators could employ flipped classrooms in teaching their courses to cover the thorough materials and to give more room for discussion and practice in their classrooms Şahin (2019). Hence, the pre-service teachers are equipped with sufficient LAL and up-to-date information and awareness of the compulsory LAL they should acquire once they start their teaching careers.

Their LAL was further enhanced through joining training related to language assessments. Moreover, through repeated practice and reflections on their language assessment practices in their classes, teachers can improve their LAL. It is necessary for EFL teachers to promote LAL development with their colleagues by collaborating, refining language assessment practices, and sharing their views on areas requiring further language assessment training. Since the present study employed the narrative inquiry method, different results are likely to occur considering the emphasis of narrative inquiry on one's experience. Hence, future researchers may address the limitation regarding the number of participants engaged in this study. In addition, future researchers can explore this issue through the use of other approaches, the use of different LAL frameworks, or the examination of the impact of experiential aspects on a larger scale.

REFERENCES

- Al-Waaili, M. Z. M. (2018). *The relationship between five Omani teachers' technology use, and their teacher cognition and instructional practices: A case study at Sultan Qaboos University*. [PhD thesis]. University of Leeds. <http://etheses.whiterose.ac.uk/20759/>
- Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Language Teaching*, 36(2), 81–109. <https://doi.org/10.1017/s0261444803001903>
- Butler, Y. G., Peng, X., & Lee, J. (2021). Young learners' voices: Towards a learner-centered approach to understanding language assessment literacy. *Language Testing*, 38(3), 429–455. <https://doi.org/10.1177/0265532221992274>
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches (Fifth)*. Sage Publications, Inc.
- Crusan, D. J., Plakans, L., & Gebril, A. (2016). Writing assessment literacy: Surveying second language teachers' knowledge, beliefs, and practices. *Assessing Writing*, 28, 43–56. <https://doi.org/10.1016/j.asw.2016.03.001>

- Cui, Y., Liu, Y., Huang, Y., & Gao, Y. (2022). Developing English teachers' language assessment literacy in an EAP reform context through test design: A case study. *System*, 109, 102866. <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., & Klinger, D. A. (2010). Assessment literacy development: Identifying gaps in teacher candidates' learning. *Assessment in Education: Principles, Policy & Practice*, 17(4), 419–438. <https://doi.org/10.1080/0969594x.2010.516643>
- Gan, L., & Lam, R. (2022). A review on Language Assessment Literacy: Trends, FOCi and Contributions. *Language Assessment Quarterly*, 19(5), 503–525. <https://doi.org/10.1080/15434303.2022.2128802>
- Giraldo, F. (2018). Language Assessment Literacy: Implications for language teachers. *Profile Issues in Teachers' Professional Development*, 20(1), 179–195. <https://doi.org/10.15446/profile.v20n1.62089>
- Giraldo, F., & Murcia, D. (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>
- Guba, E. G. (1981). Criteria for assessing the trustworthiness of naturalistic inquiries. *Educational Technology Research and Development*, 29(2). <https://doi.org/10.1007/bf02766777>
- Hutami, D. T., & Putro, N. H. P. S. (2023). Investigating teachers' language assessment literacy in the implementation of the Merdeka Belajar (Freedom of Learning) curriculum. *VELES: Voices of English Language Education Society*, 7(1), 44–56. <https://doi.org/10.29408/veles.v7i1.7624>
- Jawhar, S. S., & Subahi, A. M. (2020). The impact of specialty, sex, qualification, and experience on teachers' assessment literacy at studi higher education. *International Journal of Learning, Teaching and Educational Research*, 19(5), 200–216. <https://doi.org/10.26803/ijlter.19.5.12>
- Kim, H. K., Cho, H., Yun, H., & Shin, C. (2020). Revisiting teacher-based assessment to enhance EFL teachers' assessment literacy in South Korea. *The Journal of Asia TEFL*, 17(4), 1193–1213. <https://doi.org/10.18823/asiatefl.2020.17.4.3.1193>
- Kivunja, C., & Kuyini, A. B. (2017). Understanding and applying research paradigms in educational contexts. *International Journal of Higher Education*, 6(5), 26. <https://doi.org/10.5430/ijhe.v6n5p26>
- 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>
- Kunnan, A. J., Mathew, R., & Zhang, E. D. (2021). Language assessment literacy among teachers in Goa: Analyses of a survey questionnaire and teacher-made tests. *The Journal of Asia TEFL*, 18(3), 904–922. <https://doi.org/10.18823/asiatefl.2021.18.3.10.904>
- 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>

- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. sage.
- Malone, M. E. (2013). The essentials of assessment literacy: Contrasts between testers and users. *Language Testing*, 30(3), 329–344. <https://doi.org/10.1177/0265532213480129>
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook (Third edition)*. SAGE Publications, Inc.
- Oo, C. Z., Alonzo, D., & Asih, R. (2022). Acquisition of teacher assessment literacy by pre-service teachers: A review of practices and program designs. *Issues in Educational Research*, 32(1), 352–373.
- Restrepo-Bolívar, E. M. (2020). Monitoring pre-service teachers' language assessment literacy development through journal writing. *Malaysian Journal of ELT Research*, 17(1), 38–52.
- Richards, J. C., & Lockhart, C. (2007). *Reflective teaching in second language classrooms (Nachdr.)*. Cambridge Univ. Press.
- Şahin, S. (2019). *An Analysis of English language testing and evaluation course in English language teacher education programs in Turkey: developing language assessment literacy of pre-service efl teachers* [Ph.D. - Doctoral Program]. Middle East Technical University.
- 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>
- Sun, H., & Zhang, J. (2022). Assessment literacy of college EFL teachers in China: Status quo and mediating factors. *Studies in Educational Evaluation*, p. 74, 101157. <https://doi.org/10.1016/j.stueduc.2022.101157>
- Taylor, L. (2013). Communicating the theory, practice, and principles of language testing to test stakeholders: Some reflections. *Language Testing*, 30(3), 403–412. <https://doi.org/10.1177/0265532213480338>
- 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>
- Vogt, K., Tsagari, D., Csépes, I., Green, A., & Sifakis, N. C. (2020). Linking learners' perspectives on language assessment practices to Teachers' Assessment Literacy Enhancement (TALE): Insights from Four European Countries. *Language Assessment Quarterly*, 17(4), 410–433. <https://doi.org/10.1080/15434303.2020.1776714>
- Wang, Y., Derakhshan, A., Pan, Z., & Ghiasvand, F. (2023). Chinese EFL teachers' writing assessment feedback literacy: A scale development and validation study. *Assessing Writing*, 56, 100726. <https://doi.org/10.1016/j.asw.2023.100726>
- Xu, J., Zheng, Y., & Braund, H. (2023). Voices from L2 learners across different languages: Development and validation of a student writing assessment literacy scale. *Journal of Second Language Writing*, 60, 100993. <https://doi.org/10.1016/j.jslw.2023.100993>
- Yan, X., Cong, Z., & Fan, J. (2018). "Assessment knowledge is important, but . . .": How contextual and experiential factors mediate assessment practice and training needs of language teachers. *System*, 74, 158–168. <https://doi.org/10.1016/j.system.2018.03.003>

Zhang, C., Yan, X., & Wang, J. (2021). EFL Teachers' online assessment practices during the COVID-19 Pandemic: Changes and mediating factors. *The Asia-Pacific Education Researcher*, 30(6), 499–507. <https://doi.org/10.1007/s40299-021-00589-3>