

Research Article

Investigating Skills-Based Language Assessment Literacy of EFL Teachers in Indonesia

Wiyaka¹, Entika Fani Prastikawati^{1*}, Maria Yosephin Widarti Lestari¹, and Mahani Mokhtar²

¹English Education Department, Universitas PGRI Semarang, 50232, Indonesia

²Department of Educational Foundation and Science Social, Universiti Teknologi Malaysia, 81310, Malaysia

ORCID

Entika Fani Prastikawati: <https://orcid.org/0000-0003-4769-9915>

Abstract.

As English as a Foreign Language (EFL) teachers, they must possess a sufficient understanding of assessment-related processes. Hence, it is critical to investigate the language assessment expertise possessed by language instructors as an initial measure towards developing language assessment literacy, to identify their proficiencies and limitations in the field of language assessment. To fill the gaps, this current study examines the level of language assessment knowledge among EFL teachers in higher education schools in Indonesia. This study employed the survey method and utilized 60 adopted items, consisting of four constructs. The study was conducted among a sample of 114 EFL teachers employed in both private and public schools in Indonesia. The statistical analysis indicated that, on average, the teachers' level of language assessment knowledge is at a certain level. The findings of a one-sample t-test indicated a statistically significant difference between the scores obtained by teachers on the scale and half of the overall score. Additionally, the study revealed that the teachers demonstrated the highest level of expertise in evaluating reading skills, while their proficiency in assessing listening skills was comparatively lower. In conclusion, the current study provides several recommendations for both future research endeavors and policymakers in order to enhance the language assessment literacy of EFL instructors in evaluating individual language skills.

Keywords: skills-based language assessment, language assessment literacy, EFL teachers

1. Introduction

The integration of teaching and assessment is essential as assessment serves as an integral component within the broader framework of the learning and teaching process. Additionally, teachers are actively involved in the process of assessment and various assessment-related tasks for a significant portion of their working hours. Effective assessment methods are of paramount importance as the caliber of assessments employed serves as a requirement for the caliber of instruction and learning outcomes

Corresponding Author: Entika Fani Prastikawati; email: entikafani@upgris.ac.id

Published 12 March 2024

Publishing services provided by Knowledge E

© Wiyaka et al. This article is distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use and redistribution provided that the original author and source are credited.

Selection and Peer-review under the responsibility of the ICESRE Conference Committee.

 OPEN ACCESS

[1-3]. These assessment practices serve as valuable tools for teachers to gauge the effectiveness of their instruction [4]. By utilizing these practices, teachers can adjust the pace of their lessons, evaluate the relevance of course content, shape student learning during the teaching process, assess the efficacy of their teaching methods, and instill confidence in students for national standardized tests.

As part of their professional duties, a language instructor is entrusted with the obligation of conducting assessments. Due to the interconnected nature of the ideas underlying both teaching and assessment, both can enrich the other [5-7]. Therefore, teachers have significant responsibilities in facilitating the connection between these two notions. The crucial role of teachers in conducting assessments is emphasized, and numerous scholars in the literature [8-10] have argued that language teachers can enhance their ability to make informed decisions by acquiring knowledge in language assessment. The quality of education is significantly influenced by teachers' understanding of assessment, as evidenced by the important role they play in language assessment [11]. Consequently, it is imperative for teachers to employ assessment methodologies in order to make informed decisions, determine the most appropriate instructional approaches for students, and gain insights into the advancement of teaching and learning.

Assessment literate teachers possess a comprehensive understanding of the key components involved in the assessment process [12]. They are cognizant of the specific aspects being assessed, the underlying rationale for conducting the assessment, the most effective methods for evaluating the targeted skills and knowledge, the ability to provide high-quality exemplars of student performance, the potential pitfalls associated with the assessment, and the strategies to mitigate such issues [13]. Assessment literacy encompasses both the understanding of assessment principles and the effective implementation of these principles in the context of assessment activities. In recent times, there has been a notable emergence of a novel concept known as language assessment literacy (LAL) [14]. This concept, while originating from the broader notion of assessment literacy, has evolved into a unique domain in its own right.

The literature presents varying definitions of language evaluation literacy. According to [15], language assessment literacy refers to the level of proficiency that language teachers possess in terms of understanding testing concepts and effectively applying this knowledge to their classroom practices, particularly in relation to language assessment difficulties. According to [16], many levels of assessment literacy skills are intertwined with language-specific competences to form the knowledge foundation of language assessment. This mash-up can be seen of as language evaluation literacy

in its own right. To add, according to [17], Literacy in language assessment refers to the teacher's familiarity with and skill with a wide range of assessment-related topics and practices, such as assessment theory and practice, measurement, test design and administration, test-impact theory and practice, and assessment procedures. Therefore, as delineated in the aforementioned criteria, language assessment literacy necessitates supplementary proficiencies pertaining to language in contrast to assessment literacy. Nevertheless, the field of language assessment literacy remains in its infancy stages of development.

Some studies on LAL were done in Europe and Asian such as Hongkong, Europe, and other foreign countries. In a study conducted by [17] the objective was to examine the impact of two language assessment courses on the development of language assessment literacy among pre-service teachers in five educational institutions in Hong Kong. In the same vein, [18] conducted mixed-design research to investigate the perspectives of instructors on LAL and their specific requirements in the field of language assessment. The findings of the study indicate that the participants' perception of their level of LAL was inadequate, and they did not feel ready for assessment-related tasks. Another study conducted by [26] examined the perspectives of Turkish EFL teachers regarding in-class language evaluation and its relationship with their instructional strategies. The results indicated that teachers demonstrated a level of familiarity with fundamental concepts pertaining to classroom assessment.

The existing literature on language assessment literacy among EFL teachers primarily focuses on the assessment needs of English language teachers. This body of research highlights the inadequacy of pre-service education and the absence of sufficient professional development opportunities. Additionally, it sheds light on the self-reported assessment knowledge and practices of the participants. Hence, it is essential to investigate the language assessment expertise possessed by language teachers as an initial measure towards developing language assessment literacy, with the aim of identifying their proficiencies and limitations in the field of language assessment. The purpose of this research is to determine how well-versed EFL teachers in Indonesia are when it comes to assessing students' grasp of the English language. The study seeks to address the following research question: What is the skill-oriented level of Language Assessment Knowledge (LAK) among EFL instructors in Indonesia?

2. Method

The present study utilized the survey methodology and employed an adopted questionnaire consisting of 60 items in 3 options-Likert scale [28], which were derived from four skills. The research was carried out on a cohort of 114 EFL teachers who were engaged in both private and public educational institutions in Indonesia. The analysis of the quantitative data involved the utilization of descriptive and inferential statistics. Descriptive statistics (mean, percentage, standard deviation, etc.) and One sample t-test were utilized in presenting the result of first research question.

3. Result and Discussion

3.1. Skills-Based Language Assessment Knowledge of Indonesian EFL Teachers

The results obtained from the participants' answers are displayed in Table 1. The findings indicated that the average score of the participants in the Language Aptitude Knowledge Scale (LAKS) was 24.805. In the other words, the average number of items properly answered by the teachers was 24. Additionally, Participants' average reading comprehension test scores were the highest of every group tested in the study (6.935 out of 15). This suggests that they had a greater level of proficiency in assessing reading abilities as opposed to their proficiency in evaluating other skills. Based on the findings presented in Table 1, the participants exhibited an average score of 4.743 out of 15 when evaluating their listening abilities. Notably, this skill was identified as the area in which the participant teachers shown the lowest level of expertise.

In order to determine if the mean score is statistically and substantially lower than half of the total score, a one-sample t-test was conducted. The results are displayed in Table 2. Based on the obtained results, it was determined that there exists a statistically significant mean difference (4.76) between the average score (24.805) of the participants on the scale and half of the maximum attainable score [30], as shown by the observed data. This implies that their overall proficiency in LAK is considerably deficient.

4. Discussion

In this study, teachers demonstrated a limited understanding of how languages are evaluated. Research aimed at uncovering teachers' expertise in language evaluation

TABLE 1: Level of FL Teachers' kills-Based Language Assessment Knowledge.

ITEMS	N	True	False	Don't Know	Mean	SD
ASSESSING READING						
Requesting learners to provide a summary of the reading text serves as a method for evaluating their proficiency in reading.	114	55	54	5	.491	.503
When inquiring about multiple aspects of a reading text, each query is considered to be independent of one another.	114	33	71	10	.226	.455
The cloze test is employed as a means of evaluating the central concept or theme of a given material.	114	49	52	13	.472	.487
In the context of a reading examination, the utilization of a previously encountered text does not present an issue.	114	59	39	16	.352	.475
Including a single reading text is sufficient for the purposes of a reading examination.	114	22	83	9	.726	.447
The language complexity of the questions is lower compared to the text.	114	55	45	14	.477	.508
Spelling errors are subject to scoring penalties.	114	54	50	10	.431	.490
It is imperative to assess the level of vocabulary complexity when assessing an individual's reading abilities.	114	60	51	3	.524	.485
The inclusion of "not stated/doesn't say" options in conjunction with true/false items offers several advantages over the use of only true/false items.	114	49	45	20	.422	.490
The reliability of a reading text increases as the number of accompanying objects increases.	114	41	42	31	.360	.483
Utilizing the identical vocabulary inside the appropriate choice, as presented in the text, does not pose an issue.	114	50	51	13	.434	.492
The avoidance of simplification in reading texts is prioritized.	114	51	42	21	.376	.486
The reading materials presented in a reading examination encompass a diverse range of genres.	114	68	38	8	.610	.489
Top-down reading evaluation focuses on how well a student understands the material as a whole.	114	55	22	37	.490	.511
One issue with multiple-choice questions on a reading test is that they may include ungrammatical options.	114	61	41	12	.544	.490
READING-TOTAL	114				6.935	4.382
ASSESSING LISTENING						
There is an issue with the practice of using reading texts for listening reasons.	114	31	61	20	.281	.448
It's problematic to include unnecessary repetition [such "what I mean to say is that..."] in a listening text.	114	50	45	19	.421	.490

is rare in ELT. [18] also noted the scarcity of such research in the field of English

TABLE 1: Continued.

ITEMS	N	True	False	Don't Know	Mean	SD
ASSESSING LISTENING						
Notes can be taken from any listening source.	114	55	46	13	.415	.490
When measuring the dictation score, misspellings are disregarded.	114	18	83	13	.170	.366
Grammatical and spelling mistakes reduce possible points.	114	66	18	30	.301	.460
One method of practicing selective hearing is using a cloze listening test.	114	59	28	27	.529	.490
Integrative testing can be seen in activities like phonemic discrimination tests.	114	47	12	59	.121	.337
Note taking scoring is simple.	114	52	28	34	.240	.446
Discrete-point tests emphasize direct, surface-level understanding. .	114	41	9	64	.370	.472
Diagnosing listening abilities via dictation has no ethical concerns.	114	35	34	45	.317	.460
One reliable method of testing students' ability to listen is to provide them with a written version of the audio.	114	46	53	15	.471	.491
Distinctive-point testing is similar to dictation.	114	53	10	51	.081	.279
Listening exams rarely include inference questions designed to assess cognitive ability.	114	20	83	11	.190	.371
Having students engage in close listening by listening to lists of names or numbers is a common instructional strategy.	114	58	26	30	.240	.487
Selective listening requires students to actively seek out specific details.	114	66	38	10	.596	.498
LISTENING-TOTAL	114				4.743	3.247
ASSESSING WRITING						
Presenting learners with a choice between two choices and instructing them to write about one option guarantees scoring that is both reliable and valid.	114	64	33	17	.290	.467
Analytic scoring is employed as a means of assessing the strengths and shortcomings exhibited by learners.	114	58	36	20	.562	.524
The components of a scoring scale and the corresponding scores within each component remain consistent across varying levels of learners	114	30	69	15	.618	.483
In instances when there exists a discrepancy in the ratings assigned by the two raters, a reevaluation of the written work is conducted.	114	79	27	8	.249	.437
In the examination, students are mandated to compose responses on a minimum of two assignments, as opposed to just one task.	114	30	64	20	.277	.455
The practice of providing learners with limiting prompts or rules for writing tasks is discouraged.	114	31	69	14	.626	.472

language teaching. [15] and [11] all found that language teachers lacked expertise of

TABLE 1: Continued.

ITEMS	N	True	False	Don't Know	Mean	SD
ASSESSING WRITING						
In the realm of evaluating writing skills, it is considered a legitimate approach to present learners with a specific viewpoint and prompt them to engage in a discussion surrounding it.	114	87	14	13	.145	.330
The utilization of visual aids to assist learners in the process of writing presents a challenge.	114	10	87	14	.720	.422
Holistic scoring is used to see whether the learner is proficient or not at the end of the term.	114	53	33	25	.465	.492
Analytic scoring leads to greater reliability than holistic scoring in writing.	114	44	39	31	.384	.487
In controlled writing, learners have the chance to convey new information.	114	34	54	26	.478	.509
The most effective approach for assessing writing skills in a classroom setting is through the utilization of analytic scoring as opposed to holistic scoring	114	44	34	36	.388	.473
In process writing, the assessment of preliminary drafts does not take into account irrelevant thoughts.	114	35	60	19	.533	.460
One technique to assess a writer's proficiency is to give them a piece of reading material to write about.	114	52	41	21	.469	.490
In the subsequent stages of assessing a written work, mechanic mistakes are addressed.	114	38	61	15	.320	.465
WRITING-TOTAL	114				6.524	2.426
ASSESSING SPEAKING						
When the interlocutor fails to comprehend the learner, expressing or indicating this difficulty is a challenge.	114	66	39	9	.321	.480
Assessing the speaking skills of learners can be effectively accomplished by assigning them a single task.	114	6	101	7	.806	.334
The expression of interest by interlocutors through verbal and non-verbal signals presents a challenge.	114	26	80	8	.772	.451
Once it becomes evident that the learner is unable to attain the criterion level, the work is terminated.	114	31	67	16	.205	.450
Simultaneously employing both holistic and analytic scales is difficult to accomplish.	114	30	47	37	.420	.490
The act of reading aloud is a method employed for the evaluation of one's oral communication abilities.	114	18	79	17	.159	.339
During interlocutor-learner conversations, the instructor has the opportunity to modify the questions being posed.	114	43	58	13	.389	.483
In the context of interactive tasks, it is observed that a problem is presented by a group of learners consisting of more than two individuals.	114	30	66	18	.278	.432

language assessment, hence the results of the current study are consistent with earlier

TABLE 1: Continued.

ITEMS	N	True	False	Don't Know	Mean	SD
ASSESSING SPEAKING						
The interlocutor provides the score to the learner when they are situated within the examination room.	114	14	90	10	.779	.409
Production and understanding are both assessed concurrently during a speaking examination.	114	59	51	4	.521	.506
Requesting learners to replicate a word, phrase, or sentence as a means of evaluating their oral communication proficiency.	114	22	75	17	.217	.428
One way to gauge a student's progress in oral communication skills is through class discussion.	114	64	44	6	.560	.480
When reviewing oral presentations for an m-class, a checklist can be helpful.	114	60	37	17	.539	.485
When the focus is to assess discourse, role plays are used	114	55	34	25	.476	.504
There is no place for chance pairings in peer relationships.	114	20	71	23	.179	.366
SPEAKING-TOTAL	114				6.603	2.604
LAK-TOTAL	114				24.805	

TABLE 2: One Sample T-test.

Mean difference	df	t	p
4.76	541	-9.74	.000*

*p< .05

studies. When discussing the gravity of the situation, [7] claimed that most teachers lack appropriate information linked to language evaluation and that for most teachers, test is a four-letter word in both its literal and figurative senses. Regarding the proficiency of participant teachers in skill-based language assessment, the research findings suggest that they exhibit a higher degree of competence and knowledge when assessing reading abilities compared to their assessment of other language skills. The rationale behind the higher mean scores shown in the assessment of reading can be elucidated According to [19], most academics feel that reading is a vital talent, if not the most critical one because so much of our information and knowledge are acquired through reading. Students need to read extensively since there is so much information available to them in the classroom. Teachers naturally include assessment of reading into their practice because of the great significance placed on teaching reading skills. It is not difficult for teachers to assess their students' reading abilities because of the abundance of pre-made assessment tools.

The assessment of listening skills is seen as a problem that will resolve itself [20]. One of the outcomes of the current study is that listening is disregarded in terms of teaching and assessment, a problem also noted by [10]. Because of its complexity as a skill and the difficulties inherent in judging it, listening is undervalued.

According to the second finding, there are moderate to strong positive correlations between the various talents, and all of the items correlate positively with LAK. Based on these findings, it seems likely that enhancing teachers' expertise of assessing a single talent will also improve their LAK as a whole, as well as the other skills for which they are responsible. This discovery may cause us to view language evaluation expertise as an integrated whole with its own constituent parts. It was also shown that all forms of skill-based knowledge had either strong or moderate positive relationships with one another. Since there are strong or moderate links between the skills, it follows that as EFL teachers gain expertise in one area of evaluation, they also gain expertise in the assessment of the other skills. The results of this study provided further evidence that various forms of knowledge used in assessing a person's competence are related to one another. This is probably due to the fact that despite their differences, all skills are ultimately useful for acquiring LAK, and the reasoning behind evaluating each skill is essentially the same. When a teacher acquires more expertise in, say, the design of multiple-choice or open-ended reading activities, that expert can be applied to and built upon when teaching other subjects.

5. Conclusion

The results of this study showed that EFL teachers in the Indonesian high school education lacked the requisite general and skill-based knowledge of language assessment. Furthermore, it was observed that there was a strong positive relationship between LAK and all skills, and between all skills and one another. In the context of Indonesia, the lower scores observed in these three abilities may be attributed to inadequately designed assessment programs, insufficient provision of continuous workshops, and the inherent challenges associated with accurately evaluating these talents. The inclusion of assessment courses in pre-service education is limited to a single course offered exclusively during the final term of the English Language Teaching (ELT) curriculum. Therefore, it is challenging to comprehensively address such a wide-ranging subject within the confines of a single course, encompassing both academic and practical components. Furthermore, there is a dearth of continuous workshops that are specifically tailored for language assessment training for instructors currently employed in the field.

Finally, the findings of this study have presented numerous opportunities for subsequent investigations. This study is limited to individuals employed at senior high schools in Indonesia. The aforementioned scale has the potential to be administered to language teachers employed within ELT departments.

References

- [1] Coombe C, Vafadar H, Mohebbi H. Language assessment literacy: what do we need to learn, unlearn, and relearn? *Lang Test in Asia*. 2020 Dec;10(1):1–6.
- [2] Kremmel B, Harding L. Towards a comprehensive, empirical model of language assessment literacy across stakeholder groups: developing the language assessment literacy survey. *Lang Assess Q*. 2020 Jan;17(1):100–20.
- [3] Pastore S, Andrade HL. Teacher assessment literacy: A three-dimensional model. *Teach Teach Educ*. 2019 Aug;84:128–38.
- [4] Levi T, Inbar-Lourie O. Assessment literacy or language assessment literacy: learning from the teachers. *Lang Assess Q*. 2020;17(2):168–82.
- [5] Meijer H, Hoekstra R, Brouwer J, Strijbos JW. Unfolding collaborative learning assessment literacy: a reflection on current assessment methods in higher education. *Assess Eval High Educ*. 2020 Nov;45(8):1222–40.
- [6] Sultana N. Language assessment literacy: an uncharted area for the English language teachers in Bangladesh. *Lang Test in Asia*. 2019;9(1):1.
- [7] Firoozi T, Razavipour K, Ahmadi A. The language assessment literacy needs of Iranian EFL teachers with a focus on reformed assessment policies. *Lang Test in Asia*. 2019 Dec;9(1):1–4.
- [8] Xu Y. English language teacher assessment literacy in practice. *Second handbook of English language teaching*. 2019:517-39.
- [9] Sun Y. A Review and Discussion of In-Service EFL Teachers' Language Assessment Literacy Level in Junior High School in China. *Studies in Applied Linguistics & TESOL*. 2022;21(2):65–80.
- [10] Babaii E, Asadnia F. A long walk to language assessment literacy: EFL teachers' reflection on language assessment research and practice. *Reflective Pract*. 2019 Nov;20(6):745–60.
- [11] Prastikawati EF, Mujiyanto J, Saleh M, WuliFitriati S. Pre-service EFL Teachers' Conceptions of Assessment During Their Teaching Practicum. *KnE Social Sciences*. 2022 Dec 21:615-26.

- [12] Giraldo F. Language assessment literacy and teachers' professional development: A review of the literature. *Profile Issues in Teachers Professional Development*. 2021 Dec;23(2):265–79.
- [13] Latif MW. Exploring tertiary EFL practitioners' knowledge base component of assessment literacy: implications for teacher professional development. *Lang Test in Asia*. 2021 Dec;11(1):1–22.
- [14] Fitriyah I, Masitoh F, Widiati U. Classroom-based language assessment literacy and professional development need between novice and experienced EFL teachers. *Indonesian Journal of Applied Linguistics*. 2022 May;12(1):124–34.
- [15] Luthfiyyah R, Cahyono BY, Ivone FM, Suryati N. EFL secondary school teachers' conceptions of online assessment in emergency remote teaching: A phenomenographic study. *Studies in English Language and Education*. 2022 Sep;9(3):1041–58.
- [16] Wang X, Zuo J, Liu F, Sun Z. A Scientometric Review of Research Trends in Language Assessment Literacy. *Educ Sci (Basel)*. 2023 Feb;13(2):190.
- [17] Lam R. Language assessment training in Hong Kong: implications for language assessment literacy. *Lang Test*. 2015 Apr;32(2):169–97.
- [18] Tzagari D, Vogt K. Assessment literacy of foreign language teachers around Europe: Research, challenges and future prospects. *Papers in Language Testing and Assessment*. 2017 Jan;6(1):41–63.
- [19] Serpil ÖZ, Derin AT. Turkish EFL instructors' in-class language assessment literacy: perceptions and practices. *ELT Research Journal*. 2017;6(1):25–44.
- [20] Fadilah R, Ayudhia HY, Chairani VS, Afni F. Assessment of English Language Teaching for 21st Century: Teachers' Perspectives on Traditional and Alternative Assessment. *JADEs Journal of Academia in English Education*. 2023 Jul Education. 2023 Jul 3;4(1):108-30.