# Investigating Teachers' Language Assessment Literacy in the Implementation of the *Merdeka Belajar* (Freedom of Learning) Curriculum

**\*1Dinar Tyas Hutami, 2Nur Hidayanto Pancoro Setyo Putro**

1-2Universitas Negeri Yogyakarta, Indonesia

**\*Correspondence:**
dinartyas.2021@student.uny.ac.id

## Abstract

Teachers' language assessment literacy (LAL) is critical to the success of education, the quality of students' learning, and students' willingness to study. Yet, studies on teachers' LAL preparation to face the *Merdeka Belajar* (Freedom of Learning) curriculum are still scarce. Most LAL studies used questionnaires to assess teachers' knowledge. In contrast to that, this study utilized teachers' LAL knowledge test adapted from Al-Bahlani (2019). This study investigates teachers' self-perceived LAL consisting of competence, frequency of practice, and assessment knowledge in The Association of English Teachers Banyumas district region (MGMP) with a total of 73 participants involved by investigating teachers' self-perceived LAL consisting of competence, frequency of practice, and assessment knowledge. The current study utilized two data sources, including a questionnaire and a language assessment knowledge test to employ a quantitative methodology using MANOVA and Pearson's product-moment correlation. The results showed both strengths and weaknesses in the teachers' LAL, as well as matches and mismatches between teachers' self-perceived and shown assessment knowledge. Overall, EFL teachers in Banyumas regency are at a fair level of LAL and pre-service training in assessment was the variable with the greatest impact on teachers' LAL. Future research may also require investigating the objectives, actions, and outcomes of assessment training provided by teachers' training institutes and professional development programs.

**Keywords**: LAL, Freedom of Learning Curriculum.

## INTRODUCTION

The *Merdeka Belajar* (Freedom of Learning) curriculum is the newest ongoing curriculum implemented by the Indonesian Ministry of Education as the gatekeeper of education policy. There are radical changes that have been going on that differ from its

predecessor. One of the most striking changes is in how teachers assess students prior to the removal of the National Examination (Aulia, 2021). Formative-based assessment as a shifted assessment model of Minimum Competency Assessment from summative-based assessment (the National examination) has assumed greater significance for language teachers to have sufficient assessment concepts and procedures (Fitriyah et al., 2022). This assessment model places a larger emphasis on the learning outcomes of students in the classroom as opposed to their final test scores.

In all curricula, teachers must be able to assess their students. This demonstrates that teachers' language assessment skills are essential (Umam & Indah, 2020). However, the *Merdeka Belajar* curriculum mechanism lays a heavy focus on teachers' autonomy in designing learning and assessments, enabling teachers to become significantly more engaged and gain extensive assessment literacy. Subsequently, teachers will be required to build learning sets and assessments from scratch based on the varying needs of students in each school, as determined by their socioeconomic origins. If the teachers pose inadequate language assessment literacy, it is feared that the teacher's learning objectives would be unfocused and teachers fail to map their students' abilities which is necessary to create future plans in their classes. The greater autonomy means that the responsibility lies on teachers' solely because they do not rely on the central government to provide the set of assessment products. The language assessment literacy (henceforth LAL) level of a language teacher can be used to determine the accuracy with which an EFL teacher evaluates students' skills (Fitriyah et al., 2022).

In relation to the concern, teachers' LAL has emerged as a central topic of study and investigation in the realm of language teaching assessment (Lam, 2019; Bahtiar & Purnawarman, 2017; Fulcher, 2012). In light of the growing significance of LAL in recent years, identifying the teachers' LAL is crucial to ensure that language teacher professional development is proceeded in the right direction, thereby providing students with accurate evaluations (Sulaiman et al., 2021). In addition to the necessity of LAL, teachers have numerous assessment responsibilities (Lan & Fan, 2019). As a note, teachers also devote up to fifty percent of their professional time to classroom assessment activities to monitor their student's achievement toward learning outcomes and influence the quality of their instruction based on the quality of the assessments used (Rad, 2019; Giraldo et al., 2018; DeLuca et al., 2015; Howell, 2013). These occurrences made it abundantly evident that language teachers must enhance LAL (Anam & Putri, 2021; Latif, 2021; Widiastuti, 2021; Tsagari, 2020; Koh et al., 2017; Prasetyo, 2018). The manner in which teachers administer LAL can affect overall learning quality (Smith et al., 2014), and give a significant influence on students' achievement (Umam & Indah, 2020; Zulaiha et al., 2020), the success of teaching (Gultom, 2016), the quality of student learning (Smith et al., 2014), and students' motivation in learning (Alkharusi, 2013; Earl, 2013).

As this is a relatively recent phenomenon, empirical evidence addressing the examination of EFL teachers' LAL in the implementation of the *Merdeka Belajar* curriculum is still limited and it has yet to receive much attention from EFL Indonesian researchers (Auliya, 2022; Zulaiha, et al., 2020). The evidence is crucial as evidence of their assessment literacy level. They will be less likely to help students acquire higher levels of academic accomplishment if they lack assessment literacy (Herrera & Macías, 2015). Hence it is essential to understand how language teachers respond to this new demand and how they

possess the competencies expected to face this challenge. In addition to that, most studies reported low to moderate levels of teachers' assessment literacy (Kılıç, 2022; Massey et al., 2020; Genç, et al., 2020; Umam & Indah, 2020; Al-bahlani, 2019), the underdeveloped teachers' assessment literacy to lack of pre-service language assessment training (Bustamante, 2022; Puspawati, 2022; DeLuca & Johnson, 2017; Lam, 2015), and weak assessment knowledge results to the measures used in assessment literacy studies (Gotch & French, 2014). These aforementioned studies mostly used a survey to assess competence, practices, and knowledge. Some earlier studies utilized the Teacher Assessment Literacy Questionnaire (TALQ) (Luthfiyyah et al., 2020; Mertler, 2004; Campbell et al., 2002; Plake et al., 1993; Plake & Impara, 1993) or Assessment Literacy Inventory (ALI) (Mertler, 2009; Mertler & Campbell, 2005). To update the research, the current study utilized the LAL knowledge test adapted from Al-bahlani (2019) based on "Educational Assessment Knowledge and Skills for Teachers" (Brookhart, 2011) to assess teachers' understanding of assessment knowledge.

Furthermore, the current study aimed at measuring teachers' competence, knowledge, and practices of The Association of English Teachers (MGMP) Banyumas district region. For these reasons, the study posed to investigate (a) the extent to which Indonesian EFL teachers of the junior school are literate, (b) how teachers perceive or self-assess their assessment literacy and their relationship and correlation among EFL teachers' assessment practices and perceived skills concerning teacher sub-groups, and lastly, (c) how teachers perform in a test that measures assessment knowledge and their relationship and correlation among EFL teachers' assessment knowledge concerning teacher sub-groups.

## METHOD

### Research Context and Participants

This study surveyed 73 participants (56 females, 17 males) of the Association of English Teachers (MGMP) Banyumas district region using online and on-site surveys. According to the analysis, the majority of participants (N=50=68.5%) had received at least one assessment course before beginning their teaching careers, whereas 31.6% had never taken any assessment courses.

### Instruments

### Teachers Assessment Perceptions Questionnaire

The basis of the questionnaire is designed based on Brookhart's (2011) eleven components consisting of five assessment skills components (C1-C5) and six assessment practice (C6-C11) components. Later, these components developed into a total of 44 items. Responses on each item of teachers' assessment competence were rated from 1 (not competent) to 5 (very competent), whereas responses on each item of assessment practice were rated from 1 (never) to 5 (always).

### Teachers' LAL Knowledge Test

This study used the EFL teacher assessment knowledge test consisting of 15 multiple-choice questions adapted from Al-bahlani (2019) based on Brookhart (2011) eleven components. The Cronbach alpha coefficients were calculated and found to be 0.80 considered ideal for testing.

Table 1. Assessment Knowledge Test Questions, Its Relevance to Brookhart, and Each Mean Component

| Test Q Number | Brookhart's Statement | Mean | Test Q Number | Brookhart's Statement | Mean |
|---|---|---|---|---|---|
| All Qs | C I | 0.59 | 8 | C VII | 0.67 |
| 4 | C II | 0.32 | 9 | C VIII | 0.38 |
| 2, 6 | C III | 0.66 | 10, 11 | C IX | 0.52 |
| 1, 3, 15 | C IV | 0.51 | 12 | C X | 0.82 |
| 4, 5, 6 | C V | 0.43 | 13, 14 | C XI | 0.51 |
| 7 | C VI | 0.34 | | | |

**The procedure of Data Collection**

Quantitative data were collected throughout two distinct periods. The first phase was a survey of teachers' assessment perceptions while the second phase used the LAL knowledge test. The test was administered online using a Google form and offline by delivering a printed copy one week following the completion of the surveys. The participants were given information about the test after they did the surveys and were instructed to complete it as they started.

**Data Analysis**

Quantitative analyses of questionnaires and knowledge tests were conducted using descriptive and inferential analyses. Firstly, descriptive statistics were carried out in revealing teachers' assessment competencies, how frequently they used them, and their knowledge of language assessment. In evaluating the questionnaire at the LAL level, the researcher employed the formula of Dixon and Massey (1987). This formula divides the interval level of teachers' LAL into five groups ranging from very poor to very good. Secondly, inferential statistics through MANOVA was used to determine the significant differences in teachers' assessment competencies, assessment practices, and assessment literacy knowledge concerning different teachers' subgroups (gender, age, teachers' experience, teachers' roles, teachers' pre-service assessment training). Finally, Pearson's product moment analysis was run to see the correlation between teachers' assessment competencies, practices, and literacy knowledge of different teachers' subgroups respectively.

## FINDING AND DISCUSSION

**RQ1: The EFL Teachers' Assessment Literacy Level**

The finding emphasized that The Association of English Teachers (MGMP) Banyumas district region is at a fair level of LAL (M=13.29, SD=2.78). This finding suggests that teachers need to improve their LAL performance whilst implementing the *Merdeka Belajar* curriculum for two points, firstly, the quality of students' learning and students' learning motivation (Alkharusi, 2013).

Table 2. Banyumas EFL Teachers' Language Assessment Literacy Level

|  | Component | Mean | SD | LAL Category |
|---|---|---|---|---|
| AC | C1: Constructing/Developing and Administering Language Assessment | 12.92 | 2.6 | Fair |
|  | C2: Performance Assessment | 20.32 | 4.2 | Good |
|  | C3: Grading | 12.74 | 2.3 | Fair |
|  | C4: Communicating Assessment Results with Others | 16.48 | 3.2 | Fair |
|  | C5: Assessment Ethics | 16.51 | 3.1 | Fair |
| **AVERAGE AC** |  | **15.80** | **3.08** | **Fair** |
| AP | C6: Using Traditional Assessment Methods | 17.44 | 3.6 | Fair |
|  | C7: Using Alternative Assessment Methods | 7.21 | 2.3 | Very Poor |
|  | C8: Communicating Assessment Results to Students | 12.78 | 3.7 | Fair |
|  | C9: Assessment Standards and Criteria | 10.07 | 2 | Poor |
|  | C10: Student-Involved Assessment | 10.19 | 1.7 | Poor |
|  | C11: Non-achievement-based Grading | 9.56 | 1.9 | Poor |
| **AVERAGE AP** |  | **11.20** | **2.53** | **Poor** |
| **TOTAL AVERAGE AC + AP** |  | **13.29** | **2.78** | **Fair** |

## RQ2: Teachers' Perceived/Self-Assess in Assessment Literacy
### a. EFL Teachers' Self-Perceptions of Assessment Competences

On average, teachers regard themselves to be fairly competent in performing the *Merdeka Belajar* assessment (M=15.80, SD=3.08). This needs room for improvement due to the demands in implementing autonomous assessment which designed by teachers instead of relying on standardized framework issued by the Ministry of Education acknowledging the nature of different needs in each school as the new curriculum demand a more evidence-based assessment due to the realization that the purpose of assessment is not just about criticizing but also improving outcomes. In addition, Table 2 reported that teachers perceived themselves to be highly competent in performance assessment (C2). This demonstrates that teachers are highly believe that they are capable of evaluating students' contributions in class, determining whether or not they have understood a topic through a series of oral questions, designing performance assessment methods with well-defined objectives, establishing a rating scale for performance criteria, using that rating scale or a checklist to evaluate students' performances, and drawing valid conclusions about students' knowledge acquisition from these assessments.

Table 3. Frequencies for teachers' competence in language assessment (N = 73)

| Scale value | Category | f | 100% |
|---|---|---|---|
| 1.00-1.79 | very low competence | 2 | 1.68% |
| 1.80-2.59 | low competence | 10 | 13.74% |
| 2.60-3.39 | moderate competence | 33 | 45.80% |

| 3.40-4.19 | high competence | 23 | 31.56% |
| 4.20-5.00 | very high competence | 5 | 6.77% |

Afterward, MANOVA was run to further examine the differences between teachers' sub-groups and their language assessment competency. The results revealed no statistically significant difference for gender with F(11, 48) = 1.234, p > .05; Wilk's Λ = .780, partial η2 = .220), age with F(44, 185.6) = 1.071, p > .05; Wilk's Λ = .421, partial η2 = .195, and pre-service assessment training on teachers' language assessment competence with F (11, 48) = .796, p > .05; Wilk's Λ = .846, partial η2 = .154). This implies the lack of important roles of teachers' gender, age, and pre-service assessment course in reinforcing teacher-perceived assessment competence. This finding replicates the previous study by Al-bahlani (2019) in which gender and age have no statistical significance toward teachers' language assessment competency. However, this study differs from Al-bahlani (2019) in that the pre-service assessment training has a statistical significance toward teachers' language assessment competency. Al-bahlani (2019) stated that "The analysis revealed a statistically significant multivariate effect for pre-service assessment training on teacher language assessment competence; *F* (3.94) = 3.059, *Wilks' Lambda* = .837, *p* = .009. There were no statistically significant multivariate effects for age, gender, or in-service assessment training on teacher language assessment competence" (pp. 105-106)

In addition, this study points out the difference between Alkharusi et al., (2012, 2014) study in which gender had a significant effect on language assessment competence. Although there are distinctions between this study's participants and Alkharusi et al., (2012, 2014), this study shared one fact in common in which the majority of participants had completed at least one pre-service assessment course regardless of their degree or number of classes taught. This raises questions about the courses and professional development programs that teachers in both studies participated in. For instance, how pertinent the content was to teachers' needs and how effective it was at influencing teachers' perspectives on language assessment.

As demonstrated in Table 4, teachers' teaching experience and ages exhibited a very weak positive statistical correlation with grading and communicating assessment results to others just like the findings in Al-bahlani (2019). This shows that teachers are weakly competent at setting grades based on students' average performance, knowing what criteria should be used to set grades, knowing which student characteristics should not be utilized to set grades, and setting grades that correspond with the attainment of learning goals. In addition, teachers are skilled in assessing students' progress using portfolios, offering written feedback, communicating assessment findings to students, providing oral feedback, and communicating assessment results to parents. Other components, in contrast, have very weak negative correlations.

Table 4. Pearson Product-Moment Correlation Coefficients of Teachers' Teaching Experience and Age in Teachers' Assessment Competence

| LAL Component | Teaching Experience | Teachers' Age |
| --- | --- | --- |
| C1: Constructing/Developing and Administering Language Assessment | (r = -.021, p = .858) | (r = -.026, p = .828) |

| | | |
|---|---|---|
| C2: Performance Assessment | (r = -.040, p = .734) | (r = -.019, p = .870) |
| C3: Grading | (r = .072, p = .546) | (r = .002, p = .990) |
| C4: Communicating Assessment Results with Others | (r = .057, p = .635) | (r = .010, p = .934) |
| C5: Assessment Ethics | (r = -.052, p = .661) | (r = -.102, p = .392) |

**b. EFL Teachers' Self-Perceptions of Assessment Practices in Language Assessment**

Referring to Table 2, English teachers are at a poor level of assessment practices (M=11.20, SD=2.53). This reports what teachers do towards what they know about the *Merdeka Belajar* assessment is still lacking, hence it demands room for improvement. As shown in Table 2, the highest assessment practice done by the teachers is using traditional assessment methods (M = 17.44, SD = 3.6). This indicates that teachers most of the time use the same old way in traditional assessment methods, such as using true or false questions, multiple-choice questions, fill-in-the-blank questions, matching questions, and short essay questions. However, in the *Merdeka Belajar* curriculum, it favors more on comprehensive authentic assessment such as observation, essay, performance tasks, etc. in order to make students to be relevantly involved directly in the assessment activity. These assessments imposed by the *Merdeka Belajar* curriculum take one step ahead for the teachers to be extremely aware of students' capability and development in individual level because teachers need to be able to create various personalized assessments that are high level thinking which requires extensive understanding of which assessment tools are suitable and appropriate.

This result is in line with Al-bahlani (2019) study in which the teachers sometimes use traditional assessment methods. On the other hand, this study claimed less frequent use of alternative assessment methods (M = 7.21, SD = 2.3).

Table 5. Frequencies for teachers' practices in language assessment (N = 73)

| Scale value | Category | f | 100% |
|---|---|---|---|
| 1.00-1.79 | never practice | 2 | 3% |
| 1.80-2.59 | rarely practice | 11 | 15% |
| 2.60-3.39 | sometimes practice | 33 | 44% |
| 3.40-4.19 | most of the time practice | 12 | 16% |
| 4.20-5.00 | Always | 16 | 22% |

MANOVA shows that there is no statistical significance difference among teachers' gender F (5, 54) = 1.707, p > .05; Wilk's Λ = .864, partial η2 = .136), age F(20, 20) = 1.474, p > .05; Wilk's Λ = .605, partial η2 = .118, and pre-service assessment training F(5, 54) = 1.316, p > .05; Wilk's Λ = .891, partial η2 = .109) on teachers' language assessment practices. These findings indicate two points. Firstly, there is no significance in teachers' gender and age towards teachers' frequency of practice. Secondly, there are no important assessment courses in pre-service teachers' training programs for teachers' frequency of practice. These results on teachers' gender and age are similar to those reported by Al-bahlani (2019). Quite

opposite, one finding related to the statistical significance of pre-service assessment training on the frequency of teachers' assessment practices is found contradicted by Al-bahlani (2019). Al-bahlani (2019) stated that "The analysis revealed a statistically significant multivariate effect for pre-service assessment training on the frequency of teacher assessment practices; $F_{(3, 90)} = 2.741$, Wilks' Lambda = .824, p = .013. There were no statistically significant multivariate effects for gender, age, and in-service assessment training." (pp. 109-110)

On the other hand, Pearson's product-moment correlation resulted that this study is in line with Alkharusi et al., (2012) finding in which it has no relation between pre-service assessment training and teachers' assessment practices. Meanwhile, this study was found to be the opposite of Alkharusi et al., (2014) study in which there were statistically significant multivariate effects on the teachers' assessment practices concerning pre-service assessment training (partial $\eta^2$ = .010). As a final judgment, two studies Alkharusi et al., (2014) and Al-bahlani (2019) share the same results, thus this study highlights the importance of pre-service assessment training courses on teachers' assessment practices. Surprisingly, only one weak positive correlation was found in using alternative assessment methods (C7) (r = .214, p = .069) and teachers' age.

Table 6. Pearson Product-Moment Correlation Coefficients of Teaching Experience and Age with Teachers' Assessment Practices

| LAL Component | Teaching Experience | Teachers' Age |
|---|---|---|
| C6: Using Traditional Assessment Methods | (r = .045, p = .704) | (r = -.001, p = .994) |
| C7: Using Alternative Assessment Methods | (r = .150, p = .205) | (r = .214, p = .069) |
| C8: Communicating Assessment Results to Students | (r = .127, p = .283) | (r = .134, p = .260) |
| C9: Assessment Standards and Criteria | (r = .130, p = .273) | (r = .130, p = .272) |
| C10: Student-Involved Assessment | (r = -.149, p = .207) | (r = -.111, p = .349) |
| C11: Non-achievement-based Grading | (r = -.173, p = .144) | (r = -.186, p = .114) |

## RQ3: Teachers' Performance in Assessment Knowledge & Its Relation and Correlation to the Sub-Groups

The assessment knowledge test reported that the scores of total correct answers per test item ranged from the lowest of 0.24 (test item 15) to the highest of 0.82 (test item 12) with an average score of 0.52. This score indicates teachers' language assessment knowledge about the *Merdeka Belajar* curriculum is still very limited just like the previous studies consistently found low levels of language assessment knowledge (Xu & Brown, 2016; Alkharusi et al., 2012). Furthermore, teachers' assessment knowledge is expected to support their instruction and effectively respond to the needs and expectations of students, parents, and the school community (Herrera & Macías, 2015). The findings can be referred to their less teaching experiences and intricately linked to instructional practices as the dominance of assessment knowledge (Louw et al., 2014; Basturkmen, 2012). As a point highlight, this study employed 21 novices and 52 experienced teachers are utilized. Thus, fewer teaching

experiences are not the reason why teachers are having limited knowledge of LAL but most likely the intricately linked to instructional practices as the dominance of assessment knowledge. The findings (See Table 1) show that the highest overall performance was found in statement number ten stating *"Language teachers should be able to assist students in making sound educational decisions based on evaluation data."* It signifies teachers are competent in assisting students to analyze, plan, and track their learning, as well as comprehend the relationship between assessment and student motivation, feelings of control, and self-regulation. In contrast, the lowest score was found in statement number two in which teachers are lacking in articulating clear learning intentions. In contrast to the earlier finding, however, Al-bahlani, (2019) study found teachers were literate the most in item number two.

In line with the previous, this study analysis revealed there were no statistically significant multivariate effects for gender with $F(15, 19) = .567b$, $p > .05$; Wilk's $\Lambda = .691$, partial $\eta2 = .309$), age with $F(15, 19) = 1.091b$, $p > .05$; Wilk's $\Lambda = .537$, partial $\eta2 = .463$, educational degree with $F(30, 38) = .765b$, $p > .05$; Wilk's $\Lambda = .389$, teaching experience with $F(45, 57) = .179$, $p > .05$; Wilk's $\Lambda = .436$, pre-service assessment training with $F(15, 19) = .665b$, $p > .05$; Wilk's $\Lambda = .656$, partial $\eta2 = .344$), and teachers' role at school with $F(45, 57) = .907$, $p > .05$; Wilk's $\Lambda = .202$ to their assessment knowledge. Previous research from Alkharusi et al. (2012) reported variations in teacher assessment knowledge in which gender, pre-service assessment course, and teaching experience all played significant roles.

The output of Pearson's product-moment correlation indicates that teachers' teaching experience and their roles at school are in the interval of very weak negative correlation to weak positive correlation. Surprisingly, item 10 was the only item showing a statistical significance value with $p \leq .05$ and a weak positive correlation towards teachers' roles. Item 10 is incorporated with item 11 which emphasizes the language teachers in the Banyumas district region's ability in communicating their interpretations of assessment results and rationale to the community surrounding their students. This happened due to their teaching experiences, how long they teach, and their roles at school. However, since the correlation is reported at a weak level, it is suggested for teachers improve their capabilities in LAL for better outcomes in the future.

## CONCLUSION

The result of this study is projected to highlight important language assessment literacy competencies, practices, and knowledge in the implementation of the *Merdeka Belajar* curriculum. The consequences of this study apply to a variety of education and EFL stakeholders globally since they pertain to teachers' training programs and teaching institutions. The study found that pre-service assessment courses have a significant impact on EFL teachers' judgments of assessment competence and practice. Therefore, it is recommended that the EFL preparation program continue to give evaluation courses to language teachers to enhance their assessment literacy in the implementation of the *Merdeka Belajar* curriculum. Secondly, teachers' training programs reassess their current provisions for evaluation in language assessment. And lastly, there is an urgent need for these programs to incorporate the numerous assessment domains outlined in this study into their theoretical and practical assessment courses. In addition, the study contributed to the field of EFL and educational assessment by addressing the need for new instrument

development to measure teachers' LAL and a framework to evaluate teachers' LAL. This study's findings confirmed those of other recent research indicating that, despite improved pre-service assessment training, teachers still exhibit assessment literacy gaps. Future research may also require investigating the objectives, actions, and outcomes of assessment training provided by teachers' preparation institutes and professional development programs.

## ACKNOWLEDGMENTS

## REFERENCES

Al-bahlani, S. M. (2019). Assessment literacy: A study of EFL teachers' assessment knowledge, perspectives, and classroom behaviors. [Doctoral dissertation, The University of Arizona]. The University of Arizona. https://repository.arizona.edu/handle/10150/633240

Alkharusi, Hussain., et al. (2014). Educational assessment profile of teachers in the Sultanate of Oman. *International Education Studies*, *7*(5), 116–137. https://doi.org/10.5539/ies.v7n5p116.

Alkharusi, H. (2013). Canonical correlation models of students' perceptions of assessment tasks, motivational orientations, and learning strategies. *International Journal of Instruction, 6*(1), 21–38.

Alkharusi, H., Aldhafri, S., Alnabhani, H., & Alkalbani, M. (2012). Educational assessment attitudes, competence, knowledge, and practices: an exploratory study of muscat teachers in the sultanate of oman. *Journal of Education and Learning*, *1*(2), 217–232. https://doi.org/10.5539/jel.v1n2p217

Anam, S., & Putri, N. V. W. (2021). How literate am I about assessment: Evidence from Indonesian EFL pre-service and in-service teachers. *English Review: Journal of English Education*, *9*(2), 377–388. https://doi.org/10.25134/erjee.v9i2.4374

Aulia, Reiga W. (2021). Teachers' perception on the new policy called Merdeka Belajar; a page of lesson plan. *Fikrotuna*, *14*(2). https://doi.org/10.32806/jf.v14i02.5194

Auliya, P. K. (2022). The involvement of English teachers on minimum competency assessment (AKM). Premise: Journal of English Education, 11(1), 82. https://doi.org/10.24127/pj.v11i1.4228

Bahtiar, I., & Purnawarman, P. (2020). Investigating english teachers' comprehension in language assessment literacy (LAL). *Advances in Social Science, Education and Humanities Research*, *508.* http://creativecommons.org/licenses/by-nc/4.0/

Basturkmen, H. (2012). Review of research into the correspondence between language teachers' stated beliefs and practices. *System, 40*(2), 282–295. https://doi.org/10.1016/j.system.2012.05.001

Brookhart, S. M. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practice*, *30*(1), 3–12. https://doi.org/10.1111/j.1745-3992.2010.00195.x

Bustamante, Analynn. (2022). Language assessment literacy of middle school english teachers in Mexico. *Languages, 7*(32). https:// doi.org/10.3390/languages7010032

Campbell, C., Murphy, J. A., & Holt, J. K. (2002). *Psychometric analysis of an assessment literacy instrument: Applicability to pre service teache*rs. Paper presented at the meeting of the Mid-Western Educational Research Association.

DeLuca, Christopher., & Johnson, Sandra. (2017). Developing assessment capable teachers in this age of accountability. *Assessment in Education: Principles, Policy & Practice, 24*(2), 121-126. https://doi.org/10.1080/0969594X.2017.1297010

DeLuca, C., LaPointe-McEwan, D., & Luhanga, U. (2015). 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

Dixon, W. J., & Massey, F. J. (1987). *Pengantar Analisis Statistik/Wilfrid J. Dixon, Frank J. Massey; Penerjemah, Sri Kustamtini Samiyono; Editor, Zanzawi Soejoeti.* Rajawali

Fitriyah, I., Masitoh, F., & Widiati, U. (2022). Classroom-based language assessment literacy and professional development needed 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

Genç, E.A., Çalışkan, H., & Yüksel, A.P. (2020). Language assessment literacy level of EFL teachers: A focus on writing and speaking assessment knowledge of the teachers. *Sakarya University Journal of Education*, 274–291. https://doi.org/10.19126/suje.626156.

Giraldo, F., & Murcia, D. (2018). Language assessment literacy for pre-service teachers: course expectations from different stakeholders. *GiST Education and Learning Research Journal*, *16*(16), 56–77. https://doi.org/10.26817/16925777.425

Gotch, C. M., & French, B. F. (2014). A systematic review of assessment literacy measures. *Educational Measurement: Issues and Practice, 33*(2), 14–18. https://doi.org/10.1111/emip.12030

Gultom, Effendy. (2016). Assessment and evaluation in EFL teaching and learning. ISELT 2016 *The Fourth International Seminar on English Language and Teaching (ISELT-4): Vol 4(1). Igniting a Brighter Future of EFL Teaching and Learning in Multilingual Societies* (pp. 190-198). English Department, Faculty of Languages and Arts Universitas Negeri Padang. http://ejournal.unp.ac.id/index.php/selt/article/view/6928/5462

Herrera, L. & Macías, D. (2015). A call for language assessment literacy in the education and development of teachers of English as a foreign language. *Colomb. Appl. Linguist. J.*, *17*(2), pp. 302-312. https://doi.org/10.14483/udistrital.jour.calj.2015.2.a09

Howell, C. A. (2013). *Development and Analysis of a measurement scale for Teacher Assessment Literacy*. [Master's thesis, East Carolina University]. East Carolina University. http://hdl.handle.net/10342/1800

Kılıç, Pınar. (2022). Assessment Literacy Perceptions of Tertiary Level English Language Professionals. [Master Thesis, İhsan Doğramacı Bilkent University]. İhsan Doğramacı Bilkent University.

Koh, Kim & Burke, Lydia & Luke, Allan & Gong, Wengao & Tan, Charlene. (2017). Developing the assessment literacy of teachers in Chinese language classrooms: A focus on

assessment task design. *Language Teaching Research*, *22*(3), 264-288. https://doi.org/10.1177/1362168816684366.

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

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

Lam, R. (2019). Teacher assessment literacy: Surveying knowledge, conceptions and practices of classroom-based writing assessment in Hong Kong. *System*, *81*(1), 78–89. https://doi.org/10.1016/j.system.2019.01.006

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

Luthfiyyah, Roghibatul., Basyari, Iin Wariin., Dwiniasih, Dwiniasih. (2020). EFL secondary teachers' assessment literacy: Assessment conceptions and practices. *Journal on English as a Foreign Language*, *10*(2), 402–421.

Massey, K. D., DeLuca, C., & LaPointe-McEwan, D. (2020). Assessment literacy in college teaching: Empirical evidence on the role and effectiveness of a faculty training course. *To Improve the Academy: A Journal of Educational Development*, *39*(1). http://dx.doi.org/10.3998/tia.17063888.0039.109

Mertler, C. A. (2009). Teachers' assessment knowledge and their perceptions of the impact of classroom assessment professional development. *Improving Schools, 12*(2), 101–113. https://doi.org/10.1177/1365480209105575

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

Mertler, C. (2004). Secondary teachers' assessment literacy: Does classroom experience make a difference?. *American Secondary Education*, *33*(1), 49–64.

Plake, B. S., & Impara, J. C. (1993). Teacher assessment literacy: Development of training modules. *Annual Meeting of the National Council on Measurement in Education, Atlanta, GA, April 1993.*

Plake, B. S., Impara, J. C., & Fager, J. J. (1993). Assessment competencies of teachers: A national survey. *Educational Measurement: Issues and Practice*, *12*(4), 10–12. https://doi.org/10.1111/j.1745-3992.1993.tb00548.x

Prasetyo, A. H. (2018). Language assessment literacy development: A student-teacher's experiences in teaching practice programs. *LLT Journal: A Journal on Language and Language Teaching*, *21*(2), 199–206. https://doi.org/10.24071/llt.2018.210209

Puspawati, I. (2022). Teachers language assessment literacy: A systematic review. Indonesian EFL Journal, 8(1), 113-126. https://doi.org/10.25134/ieflj.v8i1.5613

Rad, M. R. (2019). The Impact of EFL teachers' assessment literacy on their assessment efficiency in classroom. *Britain International of Linguistics Arts and Education (BIoLAE) Journal*, *1*(1), 9–17. https://doi.org/10.33258/biolae.v1i1.14

Smith, L., Hill, M., Cowie, B., Gilmore, A. (2014). Preparing teachers to use the enabling power of assessment. In: Wyatt-Smith, C., Klenowski, V., Colbert, P. (eds), *Designing Assessment*

*for Quality Learning*. The enabling power of assessment, vol 1. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-5902-2_19

Sulaiman, Duaa., Marandi, S. Susan., & Tajik, Leila. (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

Louw, S. R., Watson, T., & Jimarkon, P. (2014). Teacher trainers' belief about feedback on teaching practice: Negotiating the tension between authoritativeness and dialogic space. *Applied Linguistics, 37*(6), 745–764. https://doi.org/10.1093/applin/amu062

Tsagari, D. (2020). Language assessment literacy. In Hidri, Sahbi (Ed), *Perspectives on language assessment literacy*. Routledge. https://doi.org/10.4324/9781003016083

Umam, A., & Indah, Yani Awalia. (2020). Exploring in-service TEYL teachers' assessment literacy: implication for continuing professional development. *JEES (Journal of English Educators Society), 5*(1), 47–51. https://doi.org/10.21070/jees.v5i1 .364

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

Widiastuti, I. (2021). Assessment and feedback practices in the EFL classroom. *REID (Research and Evaluation in Education)*, *7*(1), 13–22. https://doi.org/10.21831/reid.v7i1.37741

Xu, Y., & Brown, G. T. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, 58, 149–162. https://doi.org/10.1016/j.tate.2016.05.010

Zulaiha, Siti & Mulyono, Herri. (2020). Exploring junior high school EFL teachers' training needs of assessment literacy. *Cogent Education, 7*(1). https://doi.org/10.1080/2331186X.2020.1772943