Cakrawala Pendidikan
Jurnal Ilmiah Pendidikan
Vol. 45 No. 1, February 2026, pp. 28-42
https://jurnal.uny.ac.id/index.php/cp/issue/view/3047
DOI: https://doi.org/10.21831/cp.v45i1.89075

Bridging the gap: Unveiling assessment literacy among non-formal education tutors

Elih Sudiapermana*, Yanti Shantini, Nike Kamarubiani, Witri Dian Rafani
Universitas Pendidikan Indonesia, Indonesia
*Corresponding Author: elsud@upi.edu

ABSTRACT

Although assessment literacy is increasingly recognized as a critical component of effective teaching, empirical research addressing this competency in non-formal education contexts remains limited. This study addresses this gap by examining the level and distribution of assessment literacy among non-formal education (NFE) tutors in Indonesia and by exploring the structural validity of an assessment literacy framework adapted to non-formal learning environments. Using survey data collected from 322 NFE tutors, the study investigates variations in assessment literacy across demographic and professional characteristics and analyzes the contribution of seven conceptual dimensions to overall assessment literacy. The findings indicate that tutors' assessment literacy levels are generally modest. No significant differences are observed in relation to gender, length of teaching experience, or type of educational work experience, whereas age and educational attainment are associated with statistically significant variation. All seven dimensions demonstrate meaningful relationships with assessment literacy, with contextual responsiveness and reflective practice emerging as particularly influential, while several indicators require further refinement. By providing empirical evidence on assessment literacy in non-formal education settings, this study contributes to the refinement of existing conceptual frameworks and underscores the importance of context-sensitive professional development and institutional support in enhancing assessment practices and learning quality in non-formal education.

Keywords: assessment literacy, non-formal education, tutors

Article history

Received:	*Revised:*	*Accepted:*	*Published:*
19 August 2025	*22 September 2025*	*04 November 2025*	*01 February 2026*

Citation (APA Style): Sudiapermana, E., Shantini, Y., Kamarubiani, N., & Rafani, W. D. (2026). Bridging the gap: Unveiling assessment literacy among non-formal education tutors. *Cakrawala Pendidikan: Jurnal Ilmiah Pendidikan, 45*(1), pp. 28-42. DOI: https://doi.org/10.21831/cp.v45i1.89075

INTRODUCTION

The ability of educators to conduct assessments is a fundamental pillar in enhancing the quality of learning. Assessment not only functions as a tool for measuring student achievement but also serves as a mechanism for guiding instructional planning, monitoring learner progress, and improving teaching strategies. Robertson et al. (2024) assert that assessment literacy holds strategic potential as a solution to enhance overall instructional quality. Moreover, well-implemented assessment practices play a crucial role in identifying and addressing learning barriers, such as academic difficulties, low engagement, and inadequate learning environments (Gaikwad et al., 2023). Through systematic assessment, educators are better positioned to adapt instructional approaches to meet diverse learner needs. The findings of Han (2024) reinforce this by showing that teachers' levels of assessment literacy significantly impact student learning outcomes, indicating that assessment competence directly affects classroom effectiveness. Similarly, Khadijeh and Amir (2015) emphasize that the quality of classroom instruction is closely linked to the effectiveness of assessment practices implemented by educators, highlighting the interdependence between teaching and assessment. In line with this, DeLuca et al. (2016) argue that since the early 1990s, assessment literacy has been a central issue in teacher education

policy and research. Within the context of language education, Sultana (2019) also emphasizes that assessment literacy constitutes a crucial professional competency.

Understanding of assessment literacy has evolved in tandem with shifts in educational paradigms and the growing diversity of evaluation practices (Pastore & Andrade, 2019). Gaikwad et al. (2023) explain that assessment literacy encompasses comprehension of various assessment forms, purposes, and techniques, as well as the capacity to apply them effectively in instructional contexts. Robertson et al. (2024) add that assessment-literate teachers are capable of actively monitoring student progress, engaging in reflective practices, and developing effective strategies to enhance learning. Jeong (2013) defines assessment literacy as the ability to understand assessment goals, available assessment types, and the implications of assessment outcomes for instructional practices. Assessment literacy is not merely a technical skill; it also entails ethical and sociocultural awareness. Pastore (2023) underscores the importance of educators' awareness of the ethical consequences of their assessment decisions. Scarino (2013) similarly argues that assessment is an integral component of the learning process that fosters student reflection and self-understanding. This aligns with the view of Bailey and Carroll (2015), who note that effective assessment can enhance students' metacognitive abilities and promote active engagement in learning.

In the context of non-formal education, approaches to assessment literacy must be adapted due to its distinct characteristics compared to formal education. Non-Formal Education (NFE) is typically characterized by its more flexible and less structured nature compared to formal educational systems, which are often institutionalized and planned (Ilieva-Trichkova & Boyadjieva, 2024). Xu and Brown (2016) highlight that the learning environment significantly influences teachers' assessment competence. Non-formal education, which is characterized by flexibility and learner-centeredness, often lacks standardized assessment structures. The absence of standardized evaluation frameworks poses a major obstacle to consistent assessment implementation in non-formal settings. Assessment literacy is crucial for tutors, because they are given full authority to assess students in non-formal equivalency education, as Government Regulation No. 57/2021 in conjunction with PP.4/2022 concerning National Education Standards, Permendikbudsaintek 21/2022 concerning Assessment Standards where the graduation of students in non-formal equivalency education (Package A, B, and C programs) is determined by educators/tutors and educational units.

The diversity of learner backgrounds in non-formal education presents a unique challenge for assessment practices. Various studies indicate that assessment practices in non-formal environments can leverage participants' prior experiences and situational contexts to enhance engagement and learning outcomes (Wizel, 2021). This necessitates the use of alternative assessment methods that not only reflect holistic learning outcomes but also foster learner participation. Luthfiyyah et al. (2020) add that authentic assessments linked to real-life contexts can stimulate learning motivation. Similarly, Whitlock and Nanavati (2013) argue that performative assessments are effective in capturing higher-order thinking skills. Consequently, non-formal tutors must be equipped to design contextually relevant assessments aligned with learners' socio-cultural backgrounds. In the context of non-formal education and equivalency education, assessment is not ideal if it relies solely on written tests, is oriented toward administrative graduation, and is not relevant to the real lives of learners. Such assessment practices tend to reduce assessment to a reporting tool, rather than a means to support learning (Xu & Brown, 2016). Therefore, the ideal assessment for non-formal education tutors needs to be authentic and formative, that is, based on real-life tasks and used to provide feedback and improve learning, rather than simply generating grades (Popham, 2018). Furthermore, assessments need to be designed flexibly and transparently to accommodate the diversity of ages and backgrounds of learners, while clarifying assessment criteria. This approach aligns with the view that assessment literacy is a contextual pedagogical practice that plays a crucial role in supporting lifelong learning and inclusive non-formal education (Yan et al., 2021).

Educators with longer teaching experience tend to have practical skills in formative assessment, but do not always demonstrate a deeper conceptual understanding of modern assessment principles (Xu & Brown, 2016). Teachers and tutors with formal education in

education generally demonstrate higher levels of assessment literacy than those from non-education backgrounds, particularly in the design of instruments and the use of explicit assessment criteria (Popham, 2018). However, in non-formal education contexts, these differences are often less pronounced because assessment practices are more influenced by contextual needs and program requirements than by academic background alone. Other research suggests that younger and early-career educators tend to be more familiar with alternative and technology-based assessment approaches, while more experienced educators rely more on established, conventional assessment practices (Yan et al., 2021). These findings confirm that assessment literacy develops as a result of the interaction between individual characteristics, professional experiences, and the institutional context in which educators work, rather than as a competency that is formed in a linear or uniform manner.

However, limited access to training and resources remains a major challenge for non-formal tutors in developing assessment literacy. Pastore (2022) notes that the lack of specialized training for non-formal tutors hinders effective assessment practices. Xu and Brown (2016) also observe that the training needs of non-formal tutors have yet to become a priority in education policy. As a result, many non-formal tutors rely solely on experience and informal assessment approaches that are rarely systematized or documented. Most studies on assessment literacy to date remain focused on formal education, leading to a limited understanding of the unique needs and competencies required in non-formal settings. Gotch and French (2014) highlight that the absence of assessment frameworks tailored to the non-formal education context hampers efforts to enhance assessment literacy in this sector. Pastore (2023) also calls for the development of relevant and adaptive models of assessment literacy suited to the dynamic nature of non-formal education. Meanwhile, Xu & Brown (2016) emphasize the need for an in-depth understanding of contextual factors in designing effective training programs. The scarcity of empirical data explaining the relationship between tutor's assessment literacy and learning outcomes in non-formal education, which impedes the development of evidence-based training policies.

In response to these challenges and gaps concerning assessment literacy among tutors in Indonesian non-formal education institutions, this article aims to explore in depth the variation in assessment literacy among non-formal education tutors based on their background characteristics. Furthermore, this study focuses on examining the construct validity and measurement instruments used to assess assessment literacy among non-formal education tutors. Based on this focus, the research questions addressed in this article are as follows: 1) How does assessment literacy among NFE Tutors vary based on gender, age, educational background, and professional experience? 2) What is the underlying construct of assessment literacy among NFE Tutors, and what is the empirical validity of the instrument used to measure it?

METHOD

This study employed a quantitative approach using a survey method, aligned with the research objective of exploring the assessment literacy of NFE tutors based on their demographic characteristics and professional experiences. The quantitative approach was chosen due to its capacity to yield objective, measurable, and statistically analyzable data, allowing the researcher to develop a structured understanding of the relationships between variables (Creswell, 2012). The subjects of this study were tutors in the Equivalence Education Program at Community Learning Centers, representing a total population of 47,540 individuals across Indonesia, as well as learning facilitators ('Pamong Belajar') working within government-run non-formal education units at the district and municipal levels (such as SKB and equivalent institutions), with a total national population of 1,780 individuals (Dapodik Kemendikdasmen, 2025). The sample was obtained randomly through voluntary participation in completing a Google Form. A total of 558 responses were collected, and based on the completeness of responses, 322 were selected for data analysis.

The research instrument was a questionnaire consisting of questions about the demographic characteristics and items measuring the NFE Tutor's assessment literacy. The demographic characteristics of respondents included the variables of age (X1), educational background (X2),

teaching experience (X3), gender (X4), and experience in other educational roles (X5). The questionnaire for the NFE Tutor's assessment literacy variables was developed based on the results of analyses of existing measurement models (Alkharusi, 2011; Chapman et al., 2015; Eggen, 2004; Gaikwad et al., 2023; Gotch & French, 2014; Kaur Jaswan Singh et al., 2023; Pastore & Andrade, 2019), and adapted by considering the differences and similarities of their theoretical frameworks. The measurement instrument of PNF tutor's assessment literacy consists of seven dimensions, namely (1) Assessment Knowledge (X6), (2) Contextual Alignment (X7), (3) Assessment Practice (X8), (4) Feedback (X9), (5) Reflection (X10), (6) Student Participation (X11), and (7) Evidence-Based Decision Making (X12). The instrument is designed in the form of a closed Likert scale questionnaire with four-point choices, ranging from very positive to very negative responses, which are adjusted to the meaning of each item. The questionnaire consists of 35 items designed to assess the seven dimensions of PNF tutor's assessment literacy (see Figure 1).

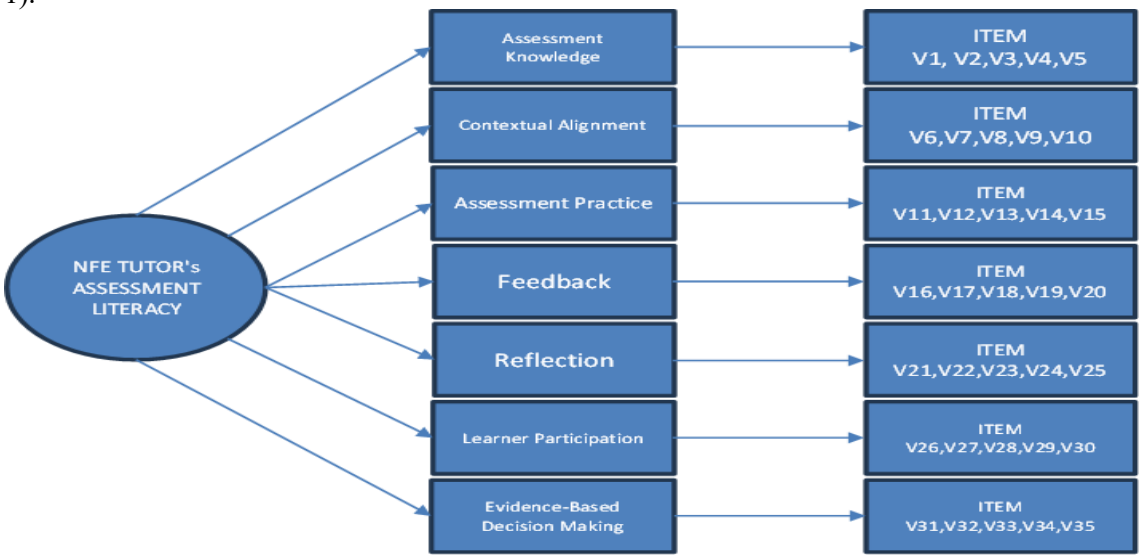

Figure 1. Constructs of NFE Tutor's Assessment Literacy

To address the research questions and test the established hypotheses, data analysis was conducted using SPSS with Descriptive, T-Test, Analysis of Variance (ANOVA), and Regression menus. The data were analyzed in three parts. First, descriptive analysis was used to describe the respondent profile and assessment literacy tendencies of NFE tutors. Second, inferential analysis aimed to test whether there were significant differences in the assessment literacy of NFE tutors based on demographic categories and professional experience. The statistical tests used included a T-test for two-group differences (female and male) and a One-Way ANOVA for more than two-group differences (between age groups, education level groups, length of experience as an NFE tutor, and work experience groups in the educational field outside of being an NFE tutor). Third, the dimensions of the NFE tutor assessment literacy measurement construct were explored using correlation analysis between item scores and the total score, as well as correlation and regression between the seven-dimension scores and the total score. Regression modeling was performed using a stepwise model based on information value criteria. To meet the analysis requirements, a normality test was conducted for the NFE tutor assessment literacy scores using a standardized residual histogram.

FINDINGS AND DISCUSSION

The results of this research offer a detailed portrait of assessment literacy levels among NFE tutors, shaped by a range of demographic factors, academic qualifications, and professional backgrounds. The quantitative data analysis identified score disparities among different groups, highlighting specific trends linked to variables such as gender, age, education level, and length of

teaching experience. In addition, the use of regression analysis and instrument validation sheds light on the multidimensional nature of assessment literacy and the overall reliability of the measurement instrument. These findings lay the groundwork for a deeper exploration of their meaning, relevance, and potential impact on efforts to strengthen assessment literacy in the non-formal education sector.

Findings

Based on the data analysis, it was found that the assessment literacy of NFE tutors tends to be low and remains far from a comprehensive level of competence. This is evident from the average scores obtained, which are considerably below the maximum possible scores. The descriptive analysis results are presented in Table 1.

Table 1. Average score of assessment literacy among tutors

Variable	N	Minimum	Maximum	Mean	Std. Deviation
Y	322	75	140	105.10	12.871
Valid N (listwise)	322	-	-	-	-

The data presented in the table above indicate that, based on a sample of 322 respondents, the average assessment literacy score was 105.10, with a score range between 75 and the maximum of 140. This suggests that the assessment literacy of tutors remains considerably below the expected maximum level (score of 140). The results of the distribution analysis show that the assessment literacy scores of NFE tutors follow a normal distribution pattern, as evidenced by the residual histogram with a mean value of 0.002 and a standard deviation of 1.001 based on 322 respondents (See Figure 2). This indicates that the assumption of normality has been met, thereby validating the suitability of the data for further analysis using predictive modeling.

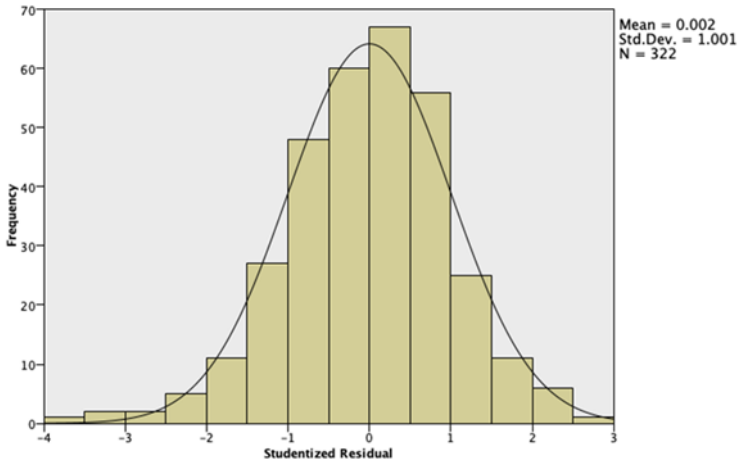

Figure 2. Residual Distribution of Assessment Literacy among NFE Tutors

Assessment Literacy Based on Various Characteristics of NFE Tutors

The results of the data analysis indicate that female Tutors (n = 127) had an average assessment literacy score of 105.24 with a standard deviation of 11.566, which is slightly higher than that of their male counterparts (n = 195), who had an average score of 105.01 and a standard deviation of 13.683. The independent t-test yielded a significance value (2-tailed) of 0.866 ($p > 0.05$), indicating that the null hypothesis, "There is no difference in assessment literacy between male and female Tutors/Learning Facilitators", cannot be rejected. This suggests that the difference in assessment literacy scores between male and female groups is not statistically significant. Therefore, although a slight difference in average scores is observed descriptively,

the analysis indicates that there is no meaningful difference in assessment literacy between male and female NFE Tutors. Detailed data are presented in Table 2.

Table 2. Results of independent samples t-test for differences in assessment literacy based on gender (Female and male)

	X4	N	Mean	Std. Deviation	Std. Error Mean
Y	1	127	105.24	11.566	1.026
	2	195	105.01	13.683	.980

The analysis results presented above indicate that female NFE (Non-Formal Education) tutors (N = 127) had an average assessment literacy score of 105.24 with a standard deviation of 11.566, which is slightly higher than that of male NFE tutors (N = 195), who scored an average of 105.01 with a standard deviation of 13.683. Therefore, although a descriptive difference in mean scores is observed, the analysis demonstrates that there is no meaningful difference in assessment literacy between male and female NFE tutors. The assessment literacy differences among NFE Tutors based on age groups are presented in the following Table 3.

Table 3. ANOVA results of differences in assessment literacy across age groups

ANOVA						
Y		Sum of Squares	df	Mean Square	F	Sig.
	Between Groups	2725.322	3	908.441	5.726	.001
	Within Groups	50453.498	318	158.659		
	Total	53178.820	321			

A one-way ANOVA test indicated that the differences in assessment literacy across age groups were statistically significant, as reflected by an F-value of 5.726 and a significance level of $p = 0.001$ ($p < 0.05$). This finding suggests that there are significant differences in assessment literacy among NFE tutors based on age group. Regarding differences in educational background among NFE tutors, the statistical analysis yielded results as shown in the following Table 4.

Table 4. ANOVA analysis of differences in assessment literacy across educational levels

ANOVA						
Y		Sum of Squares	df	Mean Square	F	Sig.
	Between Groups	3833.469	6	638.911	4.079	.001
	Within Groups	49345.351	315	156.652		
	Total	53178.820	321			

Based on statistical analysis, a significant difference was found in assessment literacy abilities among NFE Tutors with varying levels of educational attainment. These findings suggest that formal education level distinguishes the capacity for assessment literacy. Interestingly, those with lower levels of formal education, namely high school and diploma, showed higher assessment literacy scores than those with undergraduate, postgraduate, or doctoral qualifications. The analysis of differences in assessment literacy among NFE tutors based on their length of teaching experience is presented in the following Table 5.

Table 5. ANOVA results in differences in assessment literacy based on teaching experience duration

ANOVA						
Y		Sum of Squares	df	Mean Square	F	Sig.
	Between Groups	701.265	3	233.755	1.416	.238
	Within Groups	52477.555	318	165.024		
	Total	53178.820	321			

The data analysis indicates that although there are variations in the average assessment literacy scores based on the length of teaching experience as NFE tutors, these differences are not

statistically significant. The ANOVA test produced a significant value of 0.238, which exceeds the 0.05 threshold. This suggests that the duration of teaching experience does not have a significant impact on the assessment literacy level of NFE Tutors. Therefore, the length of service as a tutor in non-formal education is not a determining factor in the variation of assessment literacy competencies. Regarding differences in assessment literacy based on the diversity of professional experiences in the education field, the statistical analysis yielded results as shown in the following Table 6.

Table 6. ANOVA test for differences in assessment literacy based on types of educational work experience

	ANOVA				
Y	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1616.998	5	323.400	1.982	.081
Within Groups	51561.821	316	163.170		
Total	53178.820	321			

The data presented in the table above illustrates that there is no significant difference in assessment literacy among groups of NFE (Non-Formal Education) tutors based on variations in their work experience within the field of education. This is indicated by the ANOVA test result, which produced an F-value of 1.98 and a significant level of p = 0.081, which is greater than the conventional threshold of p = 0.05. This means that the nol hypothesis, stating that "there is no difference in assessment literacy among NFE tutors based on educational work experience", fails to be rejected. The differences in mean assessment literacy scores are not statistically significant. This finding can be interpreted to suggest that the length or type of work experience in the educational field does not significantly influence the assessment literacy skills of NFE tutors.

Constructs of NFE tutors assessment literacy

Based on the results of the regression analysis, it was found that the seven dimensions comprising the construct of assessment literacy, namely: Dimension 1: Assessment Knowledge (X6), Dimension 2: Contextual Alignment (X7), Dimension 3: Assessment Practice (X8), Dimension 4: Feedback (X9), Dimension 5: Reflection (X10), Dimension 6: Learner Participation (X11), and Dimension 7: Evidence-Based Decision Making (X12), all contributed significantly to the assessment literacy competencies of NFE Tutors, each with a significance value of p = 0.000. Among these dimensions, Contextual Alignment (X7) was identified as the most influential factor in shaping assessment literacy, as indicated by the highest importance value of 0.162. In contrast, Evidence-Based Decision Making (X12) was found to have the lowest contribution, with an important value of 0.099. Other dimensions with relatively high contributions include Reflection (X10) with an importance value of 0.160, followed by Feedback (X9) = 0.149, Learner Participation (X11) = 0.146, and Assessment Knowledge (X6) = 0.118.

These findings suggest that while all dimensions are statistically significant in shaping the overall assessment literacy of NFE tutors, their relative importance varies. This finding is reinforced by the results of the stepwise regression modeling, which demonstrate a consistent decrease in the Information Criterion value as variables were progressively included in the model (see Table 7). This indicates a hierarchical and significant contribution of the variables in explaining the variability of assessment literacy among NFE Tutors.

Overall, these results indicate a significant correlation between the seven dimensions and the level of assessment literacy among NFE Tutors, with varying weights of contribution. This supports the relevance of a multidimensional approach in measuring and enhancing assessment literacy within the context of non-formal education.

Meanwhile, the results of the item-level correlation analysis between individual questionnaire items and overall assessment literacy scores reveal that most items are positively correlated; however, there are 12 items that exhibit negative correlations with Y (assessment literacy), distributed across several dimensions. This finding suggests that not all indicators contribute positively to the assessment literacy construct. Consequently, these specific indicators

require further review and revision. The data also show that items under Dimension 7 (X12 – Evidence-Based Decision Making) contain the highest number of negatively correlated items, indicating a need for substantial refinement within this dimension. (See Figures 3).

Table 7. Stepwise regression model based on information criteria

Step	1	2	3	4	5	6	7
Information Criterion	1,229.335	1,040.811	875.198	731.371	591.955	416.499	126.501
X12_transformed	✓	✓	✓	✓	✓	✓	✓
X7_transformed		✓	✓	✓	✓	✓	✓
X10_transformed			✓	✓	✓	✓	✓
X9_transformed				✓	✓	✓	✓
X8_transformed					✓	✓	✓
X11_transformed						✓	✓
X6_transformed							✓

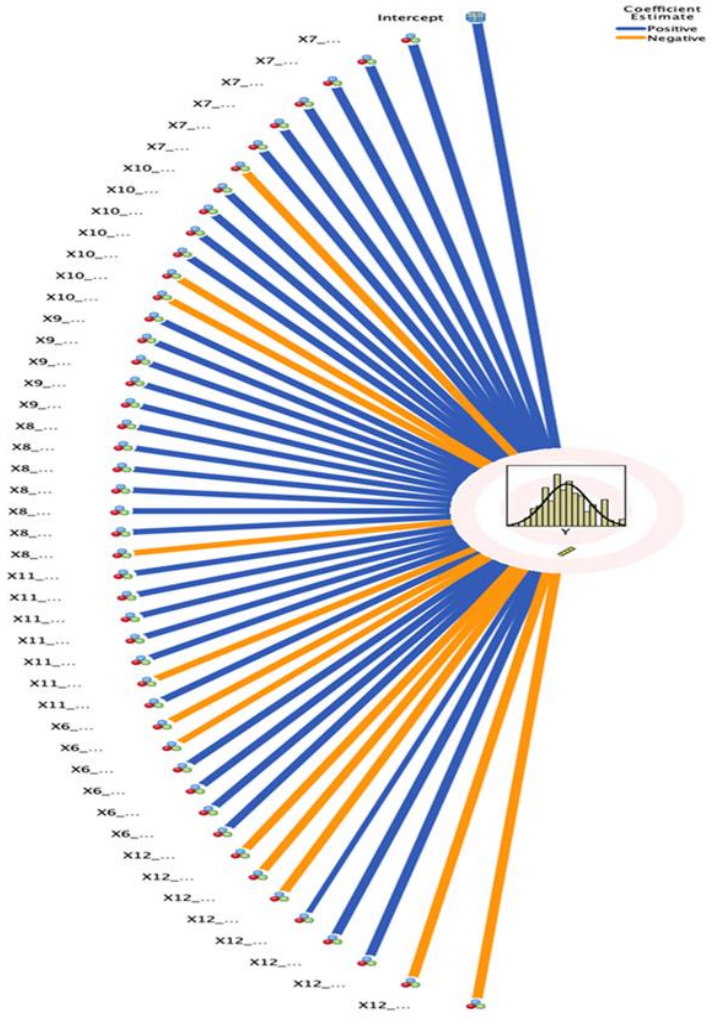

Figure 3. Visualization of Item Estimate Coefficients per Dimension on Assessment Literacy of NFE Tutors

Discussion

NFE tutors' assessment literacy based on various characteristics

The results of the hypothesis testing showed that there was no significant difference in terms of assessment literacy between female and male NFE tutors. These findings corroborate

previous research, such as that proposed by Gotch and French (2014), which found that demographic variables such as gender are not the primary determinants of assessment literacy mastery. Instead, factors such as professional experience, access to training, and institutional work culture are more influential. DeLuca et al. (2016) also noted that teachers' assessment literacy develops significantly when there is structured and ongoing training. Similarly, Koh et al. (2012) emphasized the importance of practice-based coaching and reflection to enhance teachers' understanding of the role of assessment in learning. In the context of non-formal education, this becomes even more relevant. Many tutors, especially non-formal tutors, develop assessment skills independently through field experience. As explained by Luthfiyyah et al. (2020), formal training related to assessment in non-formal education remains very limited. Therefore, improving assessment literacy should be directed at strengthening professional capacity through thematic training, mentoring, and peer learning, rather than differentiation based on personal characteristics such as gender. Although the female group scored slightly higher, this difference is more of a trend than a generalizable result. However, this trend aligns with the views of several researchers who state that female tutors tend to show greater sensitivity in responding to students' learning needs, which influences how they design, implement, and follow up on assessments. Sultana (2019) explains that affective and interpersonal dimensions are often more prominent in female teachers, which contributes to the selection of more inclusive and contextual assessment strategies. From the perspective of educational ethics, Noddings (2005) emphasizes that the ethic of care is an important foundation in learning practices, including assessment. These values are often associated with women's more participatory and reflective teaching styles. In the context of non-formal education, these findings become even more relevant. Many tutors develop their assessment skills informally through field experience. Formal training on assessment in non-formal education is still very limited. Therefore, efforts to improve assessment literacy should focus on strengthening professional capacity through thematic training, mentoring, and peer learning, rather than differentiating based on personal characteristics such as gender.

Differences in assessment literacy levels across age groups reveal an interesting pattern. Older NFE tutors achieved the highest average assessment literacy scores, followed by the youngest group, while the middle-aged group had the lowest. Increasing age, often associated with accumulated experience, has the potential to strengthen tutors' ability to understand and apply assessment principles in learning contexts. As tutors gain experience, they tend to encounter more complex and diverse learning situations. Through this process, their assessment skills are honed, both technically, such as selecting instruments and interpreting results, and pedagogically, such as adapting assessments to student characteristics. This aligns with the view of Wolterinck et al. (2024) who stated that experienced teachers are more adept at using assessment as a diagnostic tool and to support instructional decision-making. Furthermore, in non-formal education, which places a strong emphasis on flexibility and contextualization of learning, experience is a crucial factor in developing assessment intuition. Senior tutors tend to have a better understanding of students' social and cultural nuances, which is an important foundation for developing inclusive and meaningful assessments (Scarino, 2013). This perspective is further reinforced by the findings of Xu and Brown (2016), who emphasized that assessment skills develop through engagement in real-life learning practices, not solely through theoretical training. However, it is important to remember that age is not the sole indicator of assessment competency. DeLuca et al. (2016) caution that experience without the support of professional development can lead tutors to repeat old patterns that may not be appropriate to current assessment needs. Therefore, ongoing assessment training remains necessary for all age groups, including experienced tutors. These findings provide an important signal that strategies for improving assessment literacy must address the differing needs of different age groups. Younger tutors need to strengthen their conceptual foundations and practical assessment skills, while older tutors need to be guided toward refreshing their perspectives and exploring innovative assessment approaches, so that their experiences remain relevant to the demands of 21st-century learning.

The study also found a significant difference in assessment literacy levels based on the highest level of formal education attained by NFE tutors. These differences in assessment literacy performance based on educational level suggest that a higher academic degree does not

necessarily guarantee better assessment competency. This aligns with findings from several previous studies that emphasize that assessment literacy is strongly influenced by the context of teachers' training and practical experience, rather than solely by formal academic background (Hussain, 2019). Many teachers, despite their high level of education, still struggle to interpret assessment reports, indicating a gap between learned theory and actual classroom practice. Furthermore, these findings reinforce the view that assessment literacy is a multidimensional competency. This ability encompasses not only knowledge of assessment tools and techniques but also the skills to apply them appropriately in learning contexts. Chapman et al. (2015) noted in their literature review of international assessment literacy standards that even educators with advanced training often lack the practical skills to effectively apply assessment theory. Therefore, teacher training programs should integrate more practice-based exercises to bridge the gap between theory and implementation. Furthermore, higher assessment literacy scores among NFE tutors with less formal education may indicate that their practical experience is more closely connected to real-world needs, particularly in the context of flexible, student-centered non-formal learning. Although the study by Marques & Lemos (2017) focused on health literacy, their findings indicate that mastery of practical skills is often more crucial than mere theoretical knowledge. This supports the need to reconsider literacy training approaches, placing greater emphasis on strengthening the functional capacities of educators at all levels of education. The performance differences in assessment literacy based on educational background suggest that higher academic degrees do not necessarily guarantee better assessment competence. This is consistent with findings from previous studies, which emphasize that assessment literacy is strongly influenced by the context of teacher training and practical experience rather than formal academic background (Hussain, 2019; Branham et al., 2019). Many teachers, despite having advanced degrees, still struggle to interpret assessment reports—an indication of the gap between theoretical knowledge and classroom practice.

Subsequent statistical analysis revealed that although there were differences in the mean assessment literacy scores among NFE tutors based on the duration of their teaching experience, these differences were not statistically significant. This finding contrasts with the common belief that the longer a person teaches, the better their assessment skills. Herppich et al. (2014) emphasized that although experience is often associated with teaching effectiveness, empirical evidence linking tutor experience to student learning outcomes remains inconsistent. Similarly, research by Siler & VanLehn (2014) showed that tutors' assessment accuracy does not always improve with increasing experience. Karlsson (2020) added that a tutor's narrative identity within the educational environment contributes to how they support student learning. However, this contribution does not necessarily translate directly to improved academic outcomes. Furthermore, Nagraj et al. (2018) emphasized that differences in teaching experience can lead to variations in practice but do not directly improve assessment literacy skills. Based on these findings, it can be concluded that improving assessment literacy among NFE tutors should not be focused solely on the length of teaching experience. Conversely, strengthening assessment literacy can be more effective if it is directed at innovative approaches to learning strategies, such as peer tutoring and structured professional training programs (Srivastava & Rashid, 2018). Therefore, approaches to developing the capacity of non-formal education (NFE) tutors need to comprehensively consider various factors to optimally promote improved assessment literacy.

Another finding indicates that there is no difference in assessment literacy between groups of NFE tutors based on work experience in other educational fields. These findings differ from previous research on assessment practices in various educational contexts (Zulaiha & Mulyono, 2020) and highlight the importance of understanding how various dimensions of educator experience can influence assessment literacy, ultimately impacting the quality of student learning outcomes (Jenkins, 2018). Among various forms of experience, pedagogical experience demonstrated the most prominent influence, indicating a direct correlation between teaching experience and assessment skills (Al-Akbari, 2023). Furthermore, strengthening various aspects of experience can collectively optimize teacher assessment literacy (Wang et al., 2022). Meanwhile, Yazdani & Ghasedi (2021) explain that effective teaching practices typically grow

from a robust professional development process, which encourages teachers to master assessment tools and methods.

Moreover, the analysis of item-level correlation coefficients revealed that although most items showed a positive relationship with assessment literacy, a number of items displayed negative correlations. This reflects the complexity of assessment practices implemented by educators (Leirhaug & MacPhail, 2015). Therefore, ongoing professional development and reflective practices are essential to address weaknesses in specific areas of assessment literacy (Branham et al., 2019). These findings highlight the need to verify and reinforce assessment strategies through targeted and personalized training approaches to optimally improve teachers' assessment capabilities. It is also important to note that tutors with experience across various educational contexts or disciplines tend to possess richer assessment literacy, which in turn positively affects student learning outcomes. Inclusive practices such as collaborative learning and peer support further enhance this process, demonstrating that diverse experiences, combined with sustained professional engagement, fostering more effective assessment practices (Zulaiha & Mulyono, 2020; Wenner & Campbell, 2016). This is particularly relevant in non-formal education, where conventional assessment methods often face unique challenges (Mohamed et al., 2020).

In conclusion, this study supports the assertion that the diversity of tutors' experiences significantly influences the level of assessment literacy among NFE tutors. The multidimensional contributions revealed suggest that it is not merely the quantity of experience that matters, but also its quality and context. As such, professional development and reflective practice are key to addressing the challenges of non-formal education and enhancing overall learning quality.

Assessment literacy constructs and instrument quality

Based on the results of the regression analysis, it was found that all seven dimensions of the assessment literacy construct analyzed—Assessment Knowledge (X6), Contextual Alignment (X7), Assessment Practice (X8), Feedback (X9), Reflection (X10), Student Participation (X11), and Evidence-Based Decisions (X12)—significantly contributed to the assessment literacy skills of NFE tutors. This finding indicates that a small number of indicators do not fully represent the constructs being measured. This reflects the complexity of assessment practices carried out by educators (Leirhaug & MacPhail, 2015). Furthermore, according to Yusup (2018), this may lead to potential problems with instrument validity. Therefore, ongoing professional development and reflective practice are crucial to address weaknesses in specific areas of assessment literacy (Branham et al., 2019). It is also important to note that tutors with experience across multiple educational contexts or disciplines tend to have richer assessment literacy skills, which also positively impacts student learning outcomes. Inclusive practices such as collaborative learning and peer support also strengthen this process, demonstrating that diverse experiences coupled with ongoing professional engagement can foster improved assessment practices (Zulaiha & Mulyono, 2020; Wenner & Campbell, 2016). This is particularly relevant in the context of non-formal education, which presents unique challenges in implementing conventional assessment methods (Mohamed et al., 2020).

Through a modeling procedure using the forward stepwise technique, the results showed that all dimensions (X6 to X12) gradually entered the prediction model and contributed to improving the model's quality. The decrease in the Information Criterion value during the modeling process illustrates that the model became increasingly accurate and efficient in explaining variations in NFE teacher assessment literacy. This indicates that the model possesses valid and reliable characteristics, although the contribution of each dimension varies (Hapsery et al, 2019). The validity of the model depends on the cohesiveness between dimensions and internal consistency. Therefore, instrument reliability testing is also a crucial step in ensuring that the measurement tool used can produce consistent results. Overall, the analysis results indicate that despite several items not providing optimal support, the model can still be considered valid and reliable when all dimensions are considered together in the context of predicting NFE tutors' assessment literacy. The results of this research also show the need for a future research agenda to develop instruments with various methods, such as R & D (Budiastuti, et al, 2023), validity

and reliability analysis studies of instruments as conducted on teacher assessment literacy measurement instruments (Anggara, D.S & Abdilah.C, 2023; Alkharusi, H., 2011), and confirmatory factor analysis (Reise, S.P, Widaman, K.F, Pugh, R.H, 1993) to test the construct of PNF tutor assessment literacy as a result of this study's exploration.

Although this study is situated within the Indonesian nonformal education context, particularly in Paket A, Paket B, and Paket C programmes, the findings contribute to broader international discussions on assessment literacy in adult and nonformal education. International scholarship increasingly conceptualises assessment literacy not merely as an individual technical skill, but as a socially situated practice shaped by institutional norms, policy frameworks, and professional learning cultures (Xu & Brown, 2016; Pastore & Andrade, 2019). The limited variation in assessment literacy across demographic and professional characteristics observed in this study resonates with findings from adult and nonformal education settings in other countries, where assessment practices are often pragmatic, standardised, and weakly supported by systematic professional development (Yan et al., 2021). This suggests that challenges in strengthening assessment literacy among nonformal educators are not unique to Indonesia but reflect structural conditions common across lifelong learning systems internationally.

From an interdisciplinary perspective, these findings extend beyond pedagogical considerations to intersect with issues of educational governance, professional development, and social inclusion. Assessment literacy in nonformal education operates at the nexus of teaching practice, organisational learning, and public policy, influencing how learning outcomes are recognised and legitimised within broader education systems. In line with international lifelong learning literature, the study highlights that inadequate assessment capacity may constrain the role of nonformal education in supporting adult learners and second-chance education pathways (Yan et al., 2021). By situating assessment literacy within this wider interdisciplinary framework, the Indonesian case offers insights relevant for comparative research and policy discourse, reinforcing the argument that strengthening assessment literacy is a critical component of building inclusive and sustainable lifelong learning systems globally (Xu & Brown, 2016).

CONCLUSION

This study confirms that the assessment literacy of non-formal education tutors tends to be low, far from ideal. This aligns with the suspicion that attention to the non-formal education sector remains very low, including its tutors. From a gender and experience perspective (both duration and variety), there is no significant difference in assessment literacy among non-formal education tutors. This finding contrasts with research in other fields, which tends to show women outperforming men, particularly in humanitarian work. Similarly, research on work experience has found a positive correlation with increased employability. This is thought to be related to the voluntary nature of tutoring, which may result in a lack of motivation to work more professionally. Contrary to these findings, there are significant differences in assessment literacy among non-formal education tutors across age groups and educational levels. Although the pattern of differences is unique, with older and younger age groups having higher assessment literacy than middle-aged groups. Similarly, in terms of educational level, those with lower education (high school or equivalent) have higher literacy competency than those with higher education. This contrasts with various findings on educational investment, which indicate a trend toward better outcomes with higher levels of education. This is also related to the voluntary nature of being an NFE tutor. Regarding the instrument used to measure assessment literacy, regression analysis identified certain items that were negatively correlated with the assessment literacy of NFE tutors. This indicates potential issues with construct validity. However, applying a forward-stepwise regression approach revealed that all seven dimensions, from assessment knowledge to evidence-based decision-making, made statistically significant contributions to the model. While some indicators require refinement, the instrument as a whole proved valid and reliable, making it suitable for further application in research and practice. The conclusion is that assessment literacy should be recognized as a multifaceted competency shaped primarily by dynamic, experience-based processes, rather than by static demographic or academic indicators. Therefore, any strategy

aimed at strengthening assessment literacy, particularly in non-formal education contexts, should focus on experiential learning, contextual relevance, and ongoing professional growth to foster meaningful and impactful educational practice. These findings highlight the importance of ongoing, structured, and practice-oriented professional development programs to effectively improve the assessment literacy of NFE tutors. While these findings warrant further study, this research was conducted using a volunteer-based sample, making generalizations to the NFE tutor population inconclusive. Furthermore, this study tested the construct of NFE tutor assessment literacy in an exploratory manner. These two limitations should be noted for future research on NFE tutor assessment literacy.

REFERENCES

Al-Akbari, S. (2023). Developing language assessment literacy of teachers: needs, effective factors, and constraints. *Journal of Social Studies, 29*(3), 126-144. https://doi.org/10.20428/jss.v29i3.2165

Alkharusi, H. (2011). Psychometric properties of the Teacher Assessment Literacy Questionnaire for preservice teachers in Oman. *Procedia - Social and Behavioral Sciences*, *29*, 1614–1624. https://doi.org/10.1016/j.sbspro.2011.11.404

Anggara, D.S., & Abdillah, C (2023). Content validity analysis of literacy assessment instruments. *Cakrawala Pendidikan: Jurnal Ilmiah Pendidikan, 42*(2), pp.447-459. https://doi.org/10.21831/cp.v42i2.55900

Bailey, A. L., & Carroll, P. E. (2015). Assessment of English Language Learners in the Era of New Academic Content Standards. *Review of Research in Education*, *39*(1), 253–294. https://doi.org/10.3102/0091732X14556074

Branham, C. A., Blomstrom, S., Mumpower, L., Kissh, K., Wiley, J., Liu, Y., Miller, K., Bryant, Z., & Reedy, B. (2019). Digital studio tutors as partners. *International Journal for Students as Partners*, *3*(1), 140–149. https://doi.org/10.15173/ijsap.v3i1.3466

Budiastuti, E., Sugiyem, & Ahmad Puad, F. N. (2023). Developing self-assessment instruments to measure students' performance characters in making dresses using a high order thinking skills approach. *Cakrawala Pendidikan: Jurnal Ilmiah Pendidikan, 42*(1). https://doi.org/10.21831/cp.v42i1.50172

Chapman, A., Luhanga, U., & DeLuca, C. (2015). Ipsative assessment: motivation through marking progress. By Gwyneth Hughes. *British Journal of Educational Studies*, *63*(2), 246–248. https://doi.org/10.1080/00071005.2015.1035906

Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research (4th ed.)*. Pearson Education.

DeLuca, C., LaPointe-McEwan, D., & Luhanga, U. (2016). Teacher assessment literacy: a review of international standards and measures. *Educational Assessment, Evaluation and Accountability*, *28*(3), 251–272. https://doi.org/10.1007/s11092-015-9233-6

Eggen, A. B. (2004). Teacher assessment literacy beyond technicalities and intuition. *Studies in Educational Policy and Educational Philosophy*, *2004*(2), 26824. https://doi.org/10.1080/16522729.2004.11803886

Fraenkel, Jack. R., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research in education (M. Ryan, Ed.; 8th ed.)*. McGraw-Hill.

Gaikwad, S., Wadegaonkar, A., Mitra, G., & Chakravarty, D. (2023). Assessment literacy, current assessment practices and future training: reflections of teachers in higher education. *International Journal of Learning, Teaching and Educational Research*, *22*(7), 1–29. https://doi.org/10.26803/ijlter.22.7.1

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

Han, Y. (2024). A survey and analysis of assessment literacy among college foreign language teachers in China: A positive psychology perspective. *Education Reform and Development*, *6*(3). https://ojs.bbwpublisher.com/index.php/ERD

Hapsery, A., Rizki, R., & Lubis, A. (2019). Penggunaan metode stepwise pada pemodelan perencanaan track quality index (TQI) untuk kereta api semicepat Indonesia. In *MUST: Journal of Mathematics Education* (Vol. 4, Issue 1).

Herppich, S., Wittwer, J., Nückles, M., & Renkl, A. (2014). Addressing knowledge deficits in tutoring and the role of teaching experience: benefits for learning and summative assessment. *Journal of Educational Psychology, 106*(4), 934-945. https://doi.org/10.1037/a0036076

Hussain, S. (2019). A correlational study on teacher educators' assessment literacy and their students' academic achievement. *Pakistan Journal of Education, 35*(3). https://doi.org/10.30971/pje.v35i3.773

Ilieva-Trichkova, P. and Boyadjieva, P. (2024). Bounded advantages of higher education regarding young adults' participation in nonformal education. *European Journal for Research on the Education and Learning of Adults*. https://doi.org/10.3384/rela.2000-7426.4433

Jenkins, K. D. (2018). Understanding teacher beliefs and instructional decision making concerning disciplinary literacy: the case of secondary teachers in an urban school. *International Journal of Learning, Teaching and Educational Research, 17*(1), 175-196. https://doi.org/10.26803/ijlter.17.1.11

Jeong, H. (2013). Defining assessment literacy: Is it different for language testers and non-language testers? *Language Testing*, *30*(3), 345–362. https://doi.org/10.1177/0265532213480334

Karlsson, M. (2020). A question of time and place: student tutors' narrative identities in for- and non-profit contexts in sweden. *Compare: A Journal of Comparative and International Education, 51*(8), 1241-1256. https://doi.org/10.1080/03057925.2020.1835462

Kaur Jaswan Singh, H., Kaur Swaran Singh, C., Tek Ong, E., Wei Lun, W., Singh Masa Singh, T., Gopal, R., Maniam, M., Shuhaida Shukor, S., & Mulyadi, D. (2023). Developing items to measure the assessment literacy of esl teachers. In *Universitas Muhammadiyah Semarang 226 Journal of Higher Education Theory and Practice* (Vol. 23, Issue 16).

Kemendikdasmen (2025). Number of tutors. From https://ptk.datadik.kemendikdasmen.go.id/

Khadijeh, B., & Amir, R. (2015). Importance of teachers' assessment literacy. *International Journal of English Language Education*, *3*(1), 139. https://doi.org/10.5296/ijele.v3i1.6887

Koh, K. H., Tan, C., & Ng, P. T. (2012). Creating thinking schools through authentic assessment: The case in Singapore. *Educational Assessment, Evaluation and Accountability*, *24*(2), 135–149. https://doi.org/10.1007/s11092-011-9138-y

Leirhaug, P. E., & MacPhail, A. (2015). 'It's the other assessment that is the key': three Norwegian physical education teachers' engagement (or not) with assessment for learning. *Sport, Education and Society*, *20*(5), 624–640. https://doi.org/10.1080/13573322.2014.975113

Luthfiyyah, R., Basyari, I. W., & Dwiniasih, D. (2020). EFL secondary teachers' assessment literacy: Assessment conceptions and practices. *Journal on English as a Foreign Language*, *10*(2), 402–421. https://doi.org/10.23971/jefl.v10i2.2101

Marques, S. R. L., & Lemos, S. M. A. (2017). Health literacy assesment instruments: Literature review. *Audiology - Communication Research*, *22*(0). https://doi.org/10.1590/2317-6431-2016-1757

Mohamed, N. A., Samsuddin, S. F., Shaffril, H. A. M., & Bolong, J. (2020). Reading interest and reading pattern of rural library users in a low literacy rate area. *Malaysian Journal of Society and Space, 16*(4). https://doi.org/10.17576/geo-2020-1604-15

Nagraj, S., Miles, S., Bryant, P., & Holland, R. (2018). Medical students' views about having different types of problem-based learning tutors. *Medical Science Educator, 29*(1), 93-100.

Noddings, N. (2005). *The challenge to care in schools*. Teachers College Press.

Pastore, S. (2022). *Assessment literacy in the higher education context: A systematic review*. https://orcid.org/0000-0001-8598-021X

Pastore, S. (2023). Teacher assessment literacy: a systematic review. In *Frontiers in Education* (Vol. 8). Frontiers Media SA. https://doi.org/10.3389/feduc.2023.1217167

Pastore, S., & Andrade, H. L. (2019). Teacher assessment literacy: A three-dimensional model. *Teaching and Teacher Education*, *84*, 128–138. https://doi.org/10.1016/j.tate.2019.05.003

Popham, W. J. (2018). *Classroom assessment: What teachers need to know* (8th ed.). Pearson.

Reise, S.P, Widaman, K.F, Pugh, R.H (1993). Confirmatory factor analysis and item response theory: two approaches for exploring measurement invariance. *Psychological Bulletin, 114*(3), 552-566

Robertson, K. A., Hughes, K. J., & Rhind, S. M. (2024). Embedding assessment literacy can enhance graduate attribute development in a biomedical sciences curriculum. *British Journal of Biomedical Science*, *81*. https://doi.org/10.3389/bjbs.2024.12229

Scarino, A. (2013). Language assessment literacy as self-awareness: Understanding the role of interpretation in assessment and in teacher learning. *Language Testing*, *30*(3), 309–327. https://doi.org/10.1177/0265532213480128

Siler, S., VanLehn, K. (2014). Investigating microadaptation in one-to-one human tutoring. The *Journal of Experimental Education, 83*(3), 344-367. https://doi.org/10.1080/00220973.2014.907224

Srivastava, R., & Rashid, M. (2018). Who is at edge – Tutors or tutees? Academic, social and emotional elevation through peer tutoring. *Arab World English Journal (AWEJ).* Proceedings of 1st MEC TESOL conference 2018. https://doi.org/10.31235/osf.io/jph6t

Sultana, N. (2019). Language assessment literacy: an uncharted area for the English language teachers in Bangladesh. *Language Testing in Asia*, *9*(1). https://doi.org/10.1186/s40468-019-0077-8

Wang, H., Sun, W., Zhou, Y., Li, T., & Zhou, P. (2022). Teachers' assessment literacy improves teaching efficacy: a view from conservation of resources theory. *Frontiers in Psychology, 13*. https://doi.org/10.3389/fpsyg.2022.1007830

Wenner, J. A. and Campbell, T. (2016). The theoretical and empirical basis of teacher leadership. *Review of Educational Research, 87*(1), 134-171. https://doi.org/10.3102/0034654316653478

Whitlock, B., & Nanavati, J. (2013). A systematic approach to performative and authentic assessment. *Reference Services Review*, *41*(1), 32–48. https://doi.org/10.1108/00907321311300866

Wizel, M. (2021). Building bridges: bringing nonformal pedagogies into the classroom. *Education and New Developments 2021, 97*-101. https://doi.org/10.36315/2021end021

Wolterinck, C., Poortman, C., Schildkamp, K., & Visscher, A. (2024). Assessment for learning: Developing the required teacher competencies. *European Journal of Teacher Education*, *47*(4), 711–729. https://doi.org/10.1080/02619768.2022.2124912

Xu, Y., & Brown, G. T. L. (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

Yan, Z., Zhang, L. J., & Fan, J. J. (2021). Assessment literacy of teachers and teacher educators: Current status, challenges, and future directions. *Educational Measurement: Issues and Practice, 40*(3), 3–14. https://doi.org/10.1111/emip.12407

Yazdani, H., & Ghasedi, P. (2021). Efficacy of EFL teachers' assessment literacy and professional identity in boosting learners' autonomy. *Journal of Language Teaching and Research*, *12*(3), 395–403. https://doi.org/10.17507/jltr.1203.09

Yusup (2018). Uji validitas dan reliabilitas instrumen penelitian kuantitatif. Program Studi Tadris Biologi, F., & Tarbiyah dan Keguruan, F. *Januari-Juni*, *7*(1), 17–23.

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