OPEN ACCESS Crossref doi creative commons

INSTABRIGHT

INTERNATIONAL JOURNAL OF EDUCATION AND SOCIAL SCIENCES

RESEARCH ARTICLE

CLASSROOM ASSESSMENT PREFERENCE OF FOREIGN ENGLISH LANGUAGE TEACHERS IN SELECTED PRIMARY SCHOOLS IN JAKARTA, INDONESIA

MARISSA A. HUTAMARES
Batangas State University - TNEU
alzatemarissa25@gmail.com

Article History
Received: September 30, 2026
Revised: October 17, 2026
Published: October 30, 2026

ABSTRACT

This study examined the classroom language assessment preferences of Filipino and non-Filipino English teachers in selected primary schools in Jakarta, Indonesia. Using a descriptive research design, data were collected through a structured questionnaire administered to 54 teachers. Frequency, percentage, weighted mean, standard deviation, independent-samples t-test, and one-way analysis of variance were used to analyze the data. The findings revealed strong preferences for practicality, reliability, validity, and authenticity, with validity receiving the highest overall preference. Most differences in assessment preferences according to teachers' profile variables were not statistically significant. However, a significant difference was found between Filipino and non-Filipino teachers in authenticity, with Filipino teachers reporting a stronger preference. Based on these findings, the study proposed performance-based, portfolio-based, formative, rubric-based, technology-enhanced, and culturally contextualized assessment strategies, together with the integration of formative and summative assessment. The findings highlight the importance of balancing assessment quality, classroom feasibility, and meaningful language use in primary English education.

Keywords: *classroom language assessment, assessment preferences, primary English teachers, practicality, reliability, validity, authenticity*

INTRODUCTION

Language assessment is an essential part of language teaching and learning because it helps teachers evaluate students' proficiency, identify learning needs, monitor progress, and improve instruction. According to Yu et al. (2021), assessment supports both students' language development and teachers' instructional improvement.

In Indonesia, the Merdeka Curriculum emphasizes learner-centered instruction, competency development, authentic assessment, and meaningful learning experiences. Latifa et al. (2023) explained that these priorities encourage English teachers to use communicative, contextualized, and differentiated assessment approaches. However, the implementation of English assessment in primary schools remains challenging because of the subject's curriculum status and concerns about teacher preparedness. Hawanti (2014) highlighted challenges associated with the position of English as a local content subject in Indonesian primary education.

Teachers may also experience difficulties implementing authentic and formative assessments because these approaches require considerable time, effort, and assessment expertise. Consequently, some teachers continue to rely on written examinations, quizzes, and grammar-focused tests. Suryoputro and Purnawarman (2022) discussed the challenges of implementing authentic and formative assessment approaches in the Indonesian educational context.

Another concern is the gap between recommended assessment principles and actual classroom practices. Islam et al. (2021) identified this mismatch as an issue in classroom assessment, suggesting the importance of examining how assessment expectations are translated into teaching practice.

Teachers' assessment literacy is important for designing appropriate assessment tasks, providing meaningful feedback, and monitoring students' progress. Strengthening teachers' assessment knowledge and skills can help them select suitable assessment methods and use assessment results to guide instruction.

In Indonesia, strengthening teachers' professional and pedagogical competencies is also important for improving English language education. Prayogo (2022) emphasized the need for English teachers to continually develop their competencies to respond to changing educational demands.

These concerns demonstrate the importance of examining teachers' assessment preferences in relation to their professional profiles and the principles of practicality, reliability, validity, and authenticity. Therefore, this study investigated the classroom language assessment preferences of Filipino and non-Filipino primary English teachers in selected schools in Jakarta, Indonesia. The findings provide a basis for recommending practical, learner-centered, and context-appropriate assessment strategies that may support teaching practices and students' English language learning.

METHODS

The study employed a descriptive research design involving 54 Filipino and non-Filipino English teachers from selected primary schools in Jakarta, Indonesia. Data were collected through a structured questionnaire administered using Google Forms. The instrument examined teachers' classroom language assessment preferences in terms of practicality, reliability, validity, and authenticity, as well as their demographic and professional profiles.

Frequency and percentage were used to describe the respondents' profiles, while weighted mean and standard deviation summarized their assessment preferences. Independent-samples t-tests were used to compare Filipino and non-Filipino teachers, and

one-way analysis of variance was employed to examine differences across relevant profile groups. Participation depended on school approval and teachers' willingness to respond. The findings were therefore interpreted within the context of the selected schools and the study's sample of 54 respondents.

RESULTS AND DISCUSSION

1. Profile of the Respondents

1.1 Nationality

Table 2 The Profile of the Respondents in terms of Nationality

Nationality	Frequency	Percentage
Filipino	38	70.40
Non-Filipino	16	29.60
Total	54	100.00

Table 2 presents the profile of the respondents in terms of their nationality. This provides a clear overview of the representation of local, Filipino and non-Filipino teachers included in the study. Scarino (2022) which emphasized that recognizing participants' cultural and linguistic backgrounds is essential for developing inclusive and equitable instructional and assessment practices, ensuring that assessments fairly reflect students' knowledge, skills, and potential rather than privileging a single cultural or linguistic perspective.

1. 2 Educational Background

Table 3 The Profile of the Respondents in terms of Educational Background

Education	Frequency	Percentage
Bachelor's Degree	39	72.20
Master's Degree	15	27.80
Total	54	100.00

Table 3 displays the educational background of the respondents. This provides a clear overview of the respondents' level of formal education and relevant training, which may influence their knowledge, beliefs, and practices in classroom language assessment. Nearly one-third of respondents hold master's-level credentials, which has significant implications for classroom assessment. The study of Cheng & Sun (2022) supported this that teachers with advanced education often possess deeper theoretical and practical knowledge of assessment principles, such as formative and summative evaluation, differentiated assessment, and learning-oriented approaches.

1. 3 Number of Years in Teaching

Table 4 presents the profile of respondents in terms of their number of years in teaching. Studies suggest that teachers with greater experience typically exhibit higher levels of assessment literacy due to prolonged exposure to diverse learners, curricula, and assessment practices. This finding is true in the study of Scarino, 2022 which explained that these teachers are generally better able to design, implement, and interpret assessments that accurately reflect students' language proficiency while adhering to established assessment principles.

Table 4 The Profile of the Respondents in terms of the Number of Years in Teaching

Years in Teaching	Frequency	Percentage
16 years and above	12	22.20
11-15 years	19	35.20
6-10 years	12	22.20
1-5 years	11	20.40
Total	54	100.00

1.4 Number of Seminars/Training Attended

Table 5 Profile of the Respondents in terms of Number of Trainings attended on Language Assessment

No. of Seminars Attended	Frequency	Percentage
10 and above	16	29.60
7-9	8	14.80
4-6	15	27.80
Below 3	15	27.80
Total	54	100.00

Table 5 presents the number of seminars attended by the 54 respondents. This distribution reveals varied levels of engagement in continuous professional development among the respondents. Teachers' participation in seminars and training significantly influences their assessment literacy and classroom assessment practices. Brown and Abeywickrama (2019) argued that regular engagement in professional development activities enhances teachers' understanding of key assessment principles such as reliability, validity, practicality, and authenticity.

2. Classroom language assessment preferences of the respondents

Table 6 presents the assessment preferences of the respondents in terms of practicality. The respondents strongly agreed that practical assessments "provide meaningful data within reasonable limitations" and are "adaptable to different classroom sizes and settings". This result reflects teachers' need to obtain helpful information about learners' language performance without exceeding classroom constraints. The result is true to the study of Larenas et al. (2021) which explains that practicality is an essential consideration in the design and implementation of any classroom assessment, as assessments should be feasible in terms of time, resources, administration, and scoring while still achieving their intended learning objectives.

Table 6 Classroom Language Assessment Preferences of the Respondents in terms of Practicality

Classroom Language Assessment Preferences	Weighted Mean	Verbal Interpretation
1. is practical and manageable throughout the entire teaching and learning process.	3.82	Strongly Agree
2. requires minimal resources as a reference.	3.02	Agree
3. has a format that can be graded efficiently within the teacher's workload.	3.71	Strongly Agree
4. fits easily into regular classroom activities.	3.71	Strongly Agree
5. requires low logistical setup or planning.	3.18	Strongly Agree
6. involves peers or self-evaluation.	3.33	Agree
7. uses short and frequent quizzes instead of long tests administered occasionally.	3.22	Agree
8. utilizes digital tools and applications to facilitate assessment tasks.	3.40	Strongly Agree
9. provides meaningful data on language learning within reasonable limitations.	3.60	Strongly Agree
10. favors adaptable assessments that fit easily into different classroom sizes and settings	3.67	Strongly Agree
Overall Weighted Mean	3.48	Strongly Agree

The classroom language assessment preference in terms of reliability is presented in Table 7. The findings reveal that respondents strongly agree that reliability is an essential characteristic of effective language assessment. This indicates that teachers recognize the importance of using assessment tools that yield consistent and accurate results, thereby promoting fairness and credibility in evaluating students' language performance.

The findings align with Cheng and Sun's (2022) findings, which reported that EFL teachers tend to favor assessment tools with explicit criteria and standardized formats to ensure consistency across learners and assessment occasions.

Table 7 Classroom Language Assessment Preferences of the Respondents in terms of Reliability

Classroom Language Assessment Preference	Weighted Mean	Verbal Interpretation
1. generates consistent results across different sets of learners.	3.60	Strongly Agree
2. produces similar or higher results during test–retest procedures.	3.47	Strongly Agree
3. includes clear rubrics to ensure consistency in scoring tasks.	3.76	Strongly Agree
4. is standardized to provide reliable data or results.	3.58	Strongly Agree
5. yields the same results when administered multiple times to the same group of students.	3.40	Strongly Agree
6. is objective rather than subjective.	3.51	Strongly Agree
7. is reviewed and validated by language experts.	3.58	Strongly Agree
8. contains items that are clearly worded and culturally appropriate.	3.71	Strongly Agree
9. is designed with standardized and well-defined scoring rubrics.	3.71	Strongly Agree
10.undergoes pilot testing based on the curriculum design of Indonesian schools.	3.07	Agree
Overall Weighted Mean	3.54	Strongly Agree

The classroom language assessment preference in terms of validity is presented in Table 8. The data show that respondents strongly prioritize validity in classroom language assessment. Scarino (2022) argued that valid assessment practices must be context-sensitive, considering learners' linguistic and sociocultural backgrounds to ensure meaningful and equitable assessment outcomes. The findings suggest that teachers strongly value validity as a guiding principle in classroom language assessment.

Table 8 Classroom Language Assessment Preferences of the Respondents in terms of Validity

Classroom language assessment preference	Weighted Mean	Verbal Interpretation
1. directly assesses the learning objectives of the lessons.	3.82	Strongly Agree
2. reflects an accurate picture of students' linguistic capabilities.	3.69	Strongly Agree
3. measures real language skills (listening, speaking, reading, and writing) rather than memorization.	3.78	Strongly Agree
4. routinely aligns English assessments with the language curriculum or established standards.	3.73	Strongly Agree
5.avoids test items that are unrelated to the content taught.	3.73	Strongly Agree
6. provides students with a range of opportunities to demonstrate their speaking, writing, listening, and reading skills.	3.84	Strongly Agree
7. prioritizes the accurate assessment of language skills over the ease of scoring.	3.62	Strongly Agree
8. ensures that learners are being assessed on the intended learning outcomes.	3.75	Strongly Agree
9. is not limited to simple multiple-choice tests, as such formats may not fully capture learners' true language ability.	3.73	Strongly Agree
10. solicits feedback from learners to help verify the content and relevance of the assessment.	3.56	Strongly Agree
Overall Weighted Mean	3.72	Strongly Agree

Table 9 presents the classroom language assessment preferences of the respondents in terms of authenticity. It highlights the extent to which teachers value assessment tasks that reflect real-life language use and meaningful communication. The data show that the respondents strongly value authenticity in classroom language assessment. The findings are consistent with Leung (2022), who argued that authentic classroom assessments offer richer, more meaningful insights into learners' communicative abilities than traditional paper-and-pencil tests. The results indicate that teachers prefer assessment practices that move beyond conventional testing formats and focus on practical, meaningful, and contextually relevant language use.

Table 9 Classroom Language Assessment Preferences of the Respondents in terms of Authenticity

Classroom Language Assessment Preference	Weighted Mean	Verbal Interpretation
1. closely imitates real-life situations of communication.	3.67	Strongly Agree
2. requires the use of language in ways analogous to real-world contexts, such as interviewing, presenting, or role-playing.	3.64	Strongly Agree
3. reflects authentic language use and relevant communicative contexts.	3.73	Strongly Agree
4. consists of authentic tasks that serve as good indicators of language ability.	3.73	Strongly Agree
5. uses performance-based assessments, such as projects and portfolios, that resemble real-life tasks.	3.60	Strongly Agree
6. integrates activities like writing emails, giving directions, or storytelling.	3.65	Strongly Agree
7. allows students to interact with one another in communicative tasks.	3.78	Strongly Agree
8. incorporates tasks that require students to create meaning, negotiate, or solve problems through language use.	3.71	Strongly Agree
9. is task-based rather than limited to traditional paper-and-pencil tests.	3.64	Strongly Agree
10. customizes assessments to align with the real-life needs and cultural context of Indonesian learners.	3.64	Strongly Agree
Overall Weighted Mean	3.68	Strongly Agree

3. Significant Difference of Respondent's Classroom Language Assessment Preferences and Profile

Table 10 compares Filipino and non-Filipino teachers' preferences regarding practicality, reliability, validity, and authenticity in classroom language assessment. Filipino teachers obtained slightly higher mean scores across all four principles. However, the difference was statistically significant only for authenticity, indicating a significant difference between the two groups' preferences for authentic assessment. There are no statistically significant differences were found in practicality, reliability, or validity. This finding suggests that, despite their similar preferences across most assessment principles, Filipino teachers reported a significantly stronger preference for authenticity than non-Filipino teachers. As discussed by DeLuca et al. (2021), assessment practices should balance practicality with other principles to support meaningful and effective assessment. In general, the findings highlight the importance of professional development and collaborative learning in strengthening teachers' assessment literacy and promoting effective, learner-centered assessment practices.

Table 10 The Respondent's Classroom Language Assessment Preferences Group as to Nationality

Variable	Nationality	N	Mean	SD
Practicality	Filipino	38	3.51	0.06
	Non-Filipino	16	3.49	0.11
Reliability	Filipino	38	3.66	0.05
	Non-Filipino	16	3.49	0.11
Validity	Filipino	38	3.83	0.42
	Non-Filipino	16	3.69	0.09
Authenticity	Filipino	38	3.81	0.46
	Non-Filipino	16	3.61	0.99

Table 11 presents the independent-samples t-test results comparing Filipino and non-Filipino teachers' preferences for practicality, reliability, validity, and authenticity in classroom language assessment. The findings revealed no statistically significant differences in practicality, reliability, or validity, suggesting that both groups shared similar assessment preferences. This is consistent with Fulcher (2025), who emphasized the importance of

aligning assessment tasks with intended learning outcomes. However, authenticity showed a statistically significant difference with Filipino teachers reporting a stronger preference for authentic assessment than non-Filipino teachers. This suggests that nationality-related differences may be associated with how teachers prioritize real-world language use and meaningful communication. The findings indicate that nationality was associated with a significant difference in authenticity preferences, but not in the other three assessment principles.

Table 11 The Differences in Respondent's Classroom Language Assessment Preferences

Sources	Mean Difference	p-value	Decision	Interpretation
Practicality	0.02	0.838	Accept Ho	Not Significant
Reliability	0.17	0.131	Accept Ho	Not Significant
Validity	0.14	0.109	Accept Ho	Not Significant
Authenticity	0.20	0.043	Reject Ho	Significant

Table 12 compares the assessment preferences of teachers with bachelor's and master's degrees in terms of practicality, reliability, validity, and authenticity. Teachers with bachelor's degrees reported slightly higher preferences for practicality, reliability, and validity, while both groups valued authentic assessment. However, these descriptive differences alone do not establish whether educational attainment significantly influences teachers' preferences. Cheng and Sun (2022) emphasized the importance of practicality in ensuring that assessments are manageable within classroom constraints. The findings suggest that teachers from both educational groups share similar priorities in classroom language assessment, valuing assessments that are feasible to implement, consistent, aligned with learning objectives, and reflective of real-world language use. Statistical test results are needed to determine whether differences between the groups are statistically significant.

Table 12 The Respondents' Classroom Language Assessment Preferences Group According to Education

Variable	Nationality	N	Mean	SD
Practicality	Bachelor's	39	3.53	0.06
	Master's	15	3.46	0.12
Reliability	Bachelor's	39	3.65	0.05
	Master's	15	3.51	0.13
Validity	Bachelor's	39	3.82	0.04
	Master's	15	3.73	0.10
Authenticity	Bachelor's	39	3.75	0.04
	Master's	15	3.73	0.12

Table 13 presents the independent-samples t-test results comparing teachers with bachelor's and master's degrees in their preferences for practicality, reliability, validity, and authenticity. The findings revealed no statistically significant differences between the two educational groups across all four assessment principles. Both groups similarly valued manageable assessments, consistent results, alignment with learning objectives, and meaningful real-world language use.

Table 13 The differences in respondent's classroom language assessment preferences group according to Education

Sources	Mean Difference	p-value	Decision	Interpretation
Practicality	0.07	0.550	Accept Ho	Not Significant
Reliability	0.14	0.203	Accept Ho	Not Significant
Validity	0.09	0.310	Accept Ho	Not Significant
Authenticity	0.02	0.860	Accept Ho	Not Significant

These findings align with Fulcher (2020), who emphasized the importance of aligning assessment tasks with intended learning outcomes. The results suggest that educational attainment did not significantly influence teachers' classroom language assessment

preferences, indicating shared priorities in effective assessment practices regardless of degree level.

Table 14 Respondent's Classroom Language Assessment Preferences Group According to Number of Years in Teaching

Variable	Number of Years in Teaching	N	Mean	SD
Practicality	1-5	11	3.48	0.42
	6-10	12	3.68	0.23
	11-15	19	3.48	0.41
	16 and above	12	3.40	0.43
Reliability	1-5	11	3.65	0.26
	6-10	12	3.69	0.24
	11-15	19	3.55	0.43
	16 and above	12	3.57	0.39
Validity	1-5	11	3.75	0.26
	6-10	12	3.82	0.29
	11-15	19	3.73	0.34
	16 and above	12	3.88	0.26
Authenticity	1-5	11	3.72	0.35
	6-10	12	3.75	0.31
	11-15	19	3.71	0.39
	16 and above	12	3.82	0.26

Table 14 shows that teachers' preferences for classroom language assessment vary slightly according to teaching experience. Teachers with 6–10 years of experience showed the strongest preference for practicality, while those with 1–5 and 6–10 years emphasized reliability. In contrast, teachers with 16 or more years of experience placed greater emphasis on validity and authenticity. However, these descriptive differences do not establish statistical significance. For validity, high mean scores are observed across all teaching experience groups, emphasizing a shared commitment to assessments that accurately measure intended learning outcomes. The highest preference comes from teachers with 16 or more years of experience, reflecting seasoned educators' focus on alignment between learning objectives and assessment tasks. The lowest mean is recorded by the 11–15-year group, though this still indicates strong agreement.

Table 15 Respondent's Classroom Language Assessment Preferences Group According to Number of Years in Teaching

Sources	F-value	p-value	Decision	`Interpretation
Practicality	1.108	0.355	Accept Ho	Not Significant
Reliability	0.470	0.704	Accept Ho	Not Significant
Validity	0.541	0.651	Accept Ho	Not Significant
Authenticity	0.314	0.815	Accept Ho	Not Significant

Table 15 presents the one-way ANOVA results, which revealed no statistically significant differences in teachers' preferences for practicality, reliability, validity, and authenticity across teaching experience groups. Regardless of their years of service, teachers similarly valued manageable assessment tools, consistent results, alignment with learning objectives, and real-world language use. These findings align with Fulcher (2020), who emphasized the importance of aligning assessment tasks with intended learning outcomes. Overall, the results suggest that teaching experience does not significantly influence teachers' classroom language assessment preferences, reflecting shared priorities in effective language assessment.

Table 16 shows that teachers generally value practicality, reliability, validity, and authenticity regardless of the number of assessment seminars attended. Teachers who attended 4–6 seminars showed the strongest preference for practicality, while those who attended 10 or more seminars reported the highest preferences for validity and authenticity. Reliability ratings were similar across all groups. These findings align with DeLuca et al. (2023), who emphasized that professional learning could strengthen teachers' competence and confidence in implementing authentic assessment practices. The results suggest that continued professional development may help enhance teachers' assessment literacy, particularly in designing valid and authentic language assessments.

Table 16 Respondent's Classroom Language Assessment Preferences - Group According to Number of Seminars/Trainings Attended

Variable	Number of Seminars/Trainings Attended	N	Mean	SD
Practicality	Below 3	15	3.45	0.50
	4-6	15	3.58	0.30
	7-9	8	3.50	0.38
	10 and above	16	3.49	0.37
Reliability	Below 3	15	3.57	0.44
	4-6	15	3.63	0.34
	7-9	8	3.64	0.32
	10 and above	16	3.61	0.33
Validity	Below 3	15	3.73	0.38
	4-6	15	3.73	0.29
	7-9	8	3.72	0.32
	10 and above	16	3.94	0.09
Authenticity	Below 3	15	3.71	047
	4-6	15	3.68	0.28
	7-9	8	3.61	0.36
	10 and above	16	3.91	0.10

Table 17 presents the one-way ANOVA results, which revealed no statistically significant differences in teachers' preferences for practicality, reliability, validity, and authenticity based on the number of assessment-related seminars attended. Despite slight variations in mean scores, teachers across all training groups similarly valued manageable assessment tools, consistent results, alignment with learning objectives, and real-world language use. These findings align with Fulcher (2020), who emphasized the importance of validity in ensuring that assessments accurately measure intended learning outcomes. The results suggest that the number of seminars attended does not significantly influence teachers' assessment preferences, indicating shared priorities in classroom language assessment regardless of professional development exposure.

Table 17 Respondent's classroom language assessment preferences group according to Number of Seminars/Trainings Attended

Sources	F-value	p-value	Decision	Interpretation
Practicality	0.260	0.854	Accept Ho	Not Significant
Reliability	0.104	0.357	Accept Ho	Not Significant
Validity	0.147	0.106	Accept Ho	Not Significant
Authenticity	0.188	0.101	Accept Ho	Not Significant

4. Proposed Classroom Language Assessment Strategies

The proposed classroom language assessment strategies were based on teachers' preferences for validity and authenticity. These findings support performance-based, portfolio, formative, and rubric-based assessments to measure pupils' language skills and monitor progress. Menggo and Gunas (2022) highlighted the value of performance-based assessment in developing speaking proficiency. Authenticity was the only principle with a significant nationality-based difference supporting the inclusion of culturally relevant tasks. Hammond (2021) emphasized the importance of culturally responsive assessment. However, the proposed strategies are recommendations, as their effectiveness was not directly measured in this study.

CONCLUSIONS

Based on the findings of this study, several conclusions can be drawn regarding the classroom language assessment preferences of Filipino and non- Filipino English teachers:

1. The respondents varied in nationality, educational attainment, teaching experience, and participation in professional development, providing diverse perspectives on classroom language assessment.

2. The teachers demonstrated strong preferences for practicality, reliability, validity, and authenticity, highlighting the importance of these principles in classroom language assessment.

3. Assessment preferences were generally consistent across nationality, educational attainment, teaching experience, and seminar attendance. However, authenticity showed a statistically significant difference by nationality, with Filipino teachers reporting a stronger preference than non-Filipino teachers. The study did not establish the reasons for this difference.

4. Effective classroom language assessment should integrate practicality, reliability, validity, and authenticity to support fair, meaningful, and appropriate assessment practices that address learners' diverse linguistic and educational needs.

RECOMMENDATIONS

Based on the findings and conclusions of this study, the following recommendations are proposed to enhance classroom language assessment practices for Filipino and foreign English teachers.

1. The proposed classroom language assessment strategies may be evaluated and reviewed by teachers before their actual utilization to determine their appropriateness, feasibility, and responsiveness to learners' needs and classroom contexts.

2. Teachers may join regular professional development activities on modern and learner-centered language assessment. They can also work with colleagues, share effective assessment strategies, and use available resources, rubrics, and technology to support effective assessment in different classroom settings.

3. Teachers may improve their assessment skills through regular training, reflection, and professional development. They may also regularly review and adjust their assessment practices to match curriculum goals and meet the changing needs of their learners.

4. Future researchers may examine how teacher beliefs, classroom context, and student characteristics influence language assessment preferences. Further studies may also evaluate the effectiveness of assessment strategies, such as performance-based assessment and portfolios, in improving students' language outcomes across diverse EFL contexts.

REFERENCES

[1] Brown, H. D., & Abeywickrama, P. (2019). Language assessment: Principles and classroom practices (3rd ed.). Pearson Education. WorldCat

[2] Cheng, L., & Sun, Y. (2022). Teachers' classroom assessment practices in EFL contexts. Language Testing, 39(3), 345–368. https://doi.org/10.1177/02655322221077018

[3] DeLuca, C., Rickey, N., & Coombs, A. J. (2021). Exploring assessment across cultures: Teachers' approaches to assessment in the U.S., China, and Canada. Cogent Education, 8(1), Article 1921903.https://doi.org/10.1080/2331186X.2021.1921903

[3] DeLuca, C., Willis, J., Harrison, C., Cowie, B., & Coombs, A. (2023). Learning to assess: Cultivating assessment capacity in teacher education. Springer. https://doi.org/10.1007/978-981-99-6199-3

[5] Fulcher, G. (2025). Practical language testing (2nd ed.). Routledge.

[6] Hammond, Z. (2021). Culturally responsive teaching and the brain: Promoting authentic engagement and rigor among culturally and linguistically diverse students (2nd ed.). Corwin.

[7] Hawanti, S. (2014). Implementing Indonesia's English language teaching policy in primary schools: The role of teachers' knowledge and beliefs. International Journal of Pedagogies and Learning, 9(2), 162–170.

[8] Islam, S., Hasan, M. K., Sultana, S., Karim, A., & Rahman, M. M. (2021). English language assessment in Bangladesh today: Principles, practices, and problems. Language Testing in Asia, 11(1). https://doi.org/10.1186/s40468-020-00116-z

[9] Larenas, C., Diaz, A., Orellana, Y., & Villalon, M. (2021). Exploring the principles of English assessment instruments. Ensaio, 29(111), 461–483. https://doi.org/10.1590/S0104-40362021000290285

[10] Latifa, H., Ratih, K., & Maryadi. (2023). Implementing the Merdeka Curriculum in English language teaching: A study of teacher learning steps. VELES: Voices of English Language Education Society, 7(3), 640–651. https://doi.org/10.29408/veles.v7i3.24049

[11] Leung, C. (2022). Classroom-based assessment in second language education. The Modern Language Journal, 106(2), 412–429. https://doi.org/10.1111/modl.12771

[12] Menggo, S., & Gunas, T. (2022). College students' perceptions on performance-based assessment use in boosting speaking ability. International Journal of Language Education, 6(4), 423–436.

[13] Prayogo, J. A. (2022). English language teaching in Indonesia in the 21st century: What needs reinforcing and enhancing for the teachers. KnE Social Sciences. https://doi.org/10.18502/kss.v7i7.10645

[14] Scarino, A. (2022). Language teacher education in diversity – a consideration of the mediating role of languages and cultures in student learning. Language and Education, 36(2),152–169. https://doi.org/10.1080/09500782.2021.1991370

[15] Suryoputro, G., & Purnawarman, P. (2022). Implementing formative assessment in Indonesian secondary schools: Challenges and opportunities. Language Testing in Asia

[16] Yu, S., & Xu, H. (2021). Feedback and assessment in second language education during the COVID-19 and beyond. Asia-Pacific Education Researcher. https://doi.org/10.1007/s40299-021-00631-4

ACKNOWLEDGEMENT

The researcher would like to express her heartfelt gratitude and sincere appreciation to the following individuals whose support, guidance, and encouragement made this study possible:

The researcher extends sincere gratitude to Dr. Christina T. Geron, thesis adviser, Dr. Aimee M. Guia, Dr. Kristoffer Conrad M. Tejada, Dr. Alice A. Cabingan, and Dr. Ma. Leticia Jose C. Basilan, members of the panel, for their expertise, guidance, encouragement, constructive feedback, and valuable suggestions that contributed significantly to the completion and improvement of this study.

To the participants of the study, for their cooperation, participation, and time which greatly contributed to the success of this research; and

Most of all, the Almighty God, for His unending grace, wisdom, strength, and guidance throughout this academic journey. His faithfulness and love carried the researcher through every challenge and triumph. To God be all the glory.