

# Implementation of Authentic Language Assessment Tools for Speaking: A Study of Grade 5 English Teachers in Bacuag District

Jary Jaya J. Macarayo<sup>1</sup> and Annabelle N. Diaz<sup>2</sup>  
Surigao del Norte State University, Philippines<sup>1,2</sup>

**Abstract:** *This study examined the extent of implementation of authentic language assessment tools in evaluating speaking skills among Grade 5 English teachers in the Bacuag District, Surigao del Norte. Using a quantitative-descriptive design, data were collected from 19 teachers through a validated survey. Results showed that assessment practices were evident in authentic assessment characteristics and highly evident in adherence to procedural guidelines. Commonly used tools included oral recitations, role-playing, checklists, and rubrics. While demographic variables such as age, sex, and years in service showed no significant differences, a notable variation was found based on educational attainment, with graduate degree holders demonstrating stronger integration of authentic assessment characteristics. The findings highlight the need for sustained professional development to enhance teachers' capacity to use evaluative tools that promote communicative competence through real-world speaking tasks.*

**Keywords:** Authentic Assessment Tools, Speaking Skills, Performance-Based Assessment, English Language Teaching

## I. INTRODUCTION

Communicating effectively in English remains essential in the Philippine educational landscape, where English is the primary medium of instruction across subject areas. Speaking is considered the most practical and challenging to develop among the four macro skills—listening, speaking, reading, and writing. It is central to learners' academic engagement, social interaction, and communicative competence. However, recent data reflect a troubling trend in speaking proficiency. A 2023 British Council–SWS survey found that only 55% of Filipino adults can speak English, with very few rating their ability as "very good" [4]. Despite curricular emphasis on performance and fluency, many teachers continue to rely on traditional, form-focused methods of evaluation that fail to measure authentic speaking skills [5].

This study narrows its focus to the quantitative investigation of how authentic assessment tools are being implemented in the context of speaking skills assessment by Grade 5 English teachers. Authentic assessment emphasizes real-world language tasks that require learners to apply speaking skills in meaningful contexts. Quantitative data from a recent study by Zaim et al. underscore the necessity and the uneven implementation of such tools in classroom practice. While educators may understand the pedagogical value of authentic assessment, its practical use is often constrained by a lack of structured tools, training, and institutional support [20]. These limitations contribute to inconsistencies in performance-based assessments, especially when evaluating spontaneous, learner-centered speaking tasks.

This study seeks to quantitatively assess the extent to which authentic language assessment tools are implemented by Grade 5 English teachers in the Bacuag District, Surigao del Norte. It focuses on identifying the specific tools teachers employ in assessing speaking skills, evaluating the degree to which these practices align with established characteristics and guidelines of authentic assessment, and determining whether variations in implementation are significantly associated with demographic factors such as age, sex, highest educational attainment, and years of teaching experience. This study focuses on measurable trends through validated survey data and provides a statistical basis for understanding classroom-level assessment behaviors. The findings seek to offer data-driven insights for teacher training, policy



development, and resource allocation, particularly in under-resourced school contexts. Ultimately, the study supports efforts to bridge the gap between curriculum expectations and actual classroom assessment practices through the lens of quantitative evidence.

## **II. METHODOLOGY**

This study adopted a quantitative-descriptive research design to systematically examine the extent to which Grade 5 English teachers implemented authentic language assessment tools in evaluating learners' speaking skills within the public elementary schools of Bacuag District, Surigao del Norte. The design was selected to obtain measurable data on assessment practices and to analyze whether implementation levels varied according to specific teacher characteristics. The respondents included 19 Grade 5 English language teachers selected through purposive sampling based on their teaching assignments, availability, and willingness to participate in the study. The research setting comprised 11 public elementary schools in both central and remote barangays within the district, allowing for representation of a range of teaching contexts.

The researcher used a survey questionnaire developed specifically for the study to collect data. The instrument consisted of two main parts. The first part gathered the demographic profile of respondents, including variables such as age, sex, highest educational attainment, and number of years in service. The second part featured 4-point Likert-scale items that measured the extent to which authentic assessment practices were implemented in their classrooms. These items were clustered into two core dimensions: (1) characteristics of authentic assessment (e.g., real-world relevance, focus on performance, learner engagement), and (2) procedural guidelines (e.g., alignment with competencies, use of rubrics, checklists, and rating scales).

The instrument was developed through an extensive review of relevant literature and educational policies and was anchored conceptually on Wiggins' Theory of Authentic Assessment [20]. To ensure content validity, the questionnaire underwent rigorous evaluation by a panel composed of the researcher's thesis adviser, three research panel members, and a professional statistician. Their feedback informed item refinement and alignment with the research objectives. Additionally, the researcher conducted a pilot test among English teachers from a neighboring district who were not part of the study. The instrument's reliability was statistically verified using Cronbach's alpha, which yielded an acceptable level of internal consistency.

Upon securing approval from the Schools Division Superintendent and the respective school heads, the researcher personally administered the questionnaires to ensure proper item clarification and accuracy of responses. All ethical considerations were upheld, including informed consent, voluntary participation, confidentiality, and the right to withdraw.

After data collection, the researcher organized the responses and subjected them to descriptive and inferential statistical analyses. Frequency and percentage were used to describe the demographic profile of the respondents. Mean and standard deviation were employed to interpret the extent of implementation of authentic assessment practices. To determine whether significant differences existed in assessment implementation when grouped according to teacher profiles, the researcher used the Mann–Whitney U test for comparisons involving two independent groups (e.g., male vs. female; bachelor's vs. master's degree holders), and the Kruskal–Wallis H test for comparisons among three or more groups (e.g., age ranges or years in service). All inferential tests were conducted at the 0.05 level of significance.

These data analysis procedures enabled a comprehensive evaluation of the degree and variability of authentic assessment implementation across different instructional contexts and teacher demographics, contributing to a deeper understanding of the current landscape of speaking assessment practices in the Philippine elementary education setting.

## **III. LITERATURE REVIEW**

Authentic assessment has emerged as a critical framework in language education, particularly in evaluating speaking skills. Unlike traditional assessments emphasizing discrete-point grammar and vocabulary tests, authentic assessment prioritizes meaningful, real-world tasks that mirror actual communicative situations. Espinosa and López [8] assert that tasks—like storytelling, interviews, and oral presentations—require learners to activate multiple language domains in



real-time, thereby providing a more comprehensive view of communicative competence. These practices are especially relevant in the Philippine context, where English serves as a medium of instruction. However, learners often struggle to demonstrate proficiency in speaking due to outdated assessment practices [5].

The theoretical grounding of authentic assessment draws from constructivist and sociocultural learning theories, which highlight the importance of situated learning and active engagement. Firdausi [9] notes that learners develop language more effectively when assessment activities are embedded in social and contextual realities. Alam et al. [2] further argue that authentic tasks improve linguistic accuracy and promote reflection and self-monitoring, key components of long-term language development. These perspectives align with the Department of Education's K to 12 framework, which advocates for performance-based evaluations to enhance 21st-century skills such as collaboration, communication, and critical thinking [14].

Empirical research supports authentic assessment in promoting student motivation, fluency, and confidence. According to Dizon [7], performance-based and video-assisted speaking assessments improved learner self-assurance and reduced language anxiety. Similarly, Sundari and Sabarun [17] found that students were more engaged in speaking tasks when presented as meaningful roles or collaborative projects. Permatasari and Nurjati [13] emphasized that digital storytelling and project-based assessments also created inclusive spaces encouraging participation, especially among shy learners. These findings are echoed by Gao et al. [10], who explored using personalized digital tools that support speaking practice while enhancing learner autonomy and privacy.

Regarding assessment tools, a range of strategies is employed to measure speaking proficiency authentically. Analytic rubrics, which provide explicit descriptors for fluency, content, pronunciation, and coherence, are widely endorsed for their transparency and instructional value [19]. Alam et al. [2] stress that the effectiveness of rubrics hinges on teacher training and clarity of criteria. E-portfolios have also gained traction as tools that allow learners to curate and reflect on their progress over time. Chionidou-Moskofoglou et al. [6] highlight their benefits in fostering metacognitive awareness and accountability. Peer and self-assessment methods are similarly encouraged, with research showing they build a culture of shared responsibility and collaboration in classrooms [14].

Technological innovations have expanded the possibilities for authentic speaking assessment. Platforms like Flipgrid and Seesaw enable asynchronous performance tasks that reduce learner anxiety while allowing repeated practice [7]. As investigated by Lu et al. [12], automated assessment tools offer scalable and objective scoring mechanisms, though they must be used with caution in multilingual settings. Project-based approaches, including role-plays and video journals, have also proven effective in fostering creativity, narrative fluency, and cross-disciplinary learning [13].

Despite these benefits, numerous challenges hinder the widespread implementation of authentic assessment, particularly in rural or resource-limited settings. Time constraints and large class sizes are persistent issues, with teachers reporting difficulty administering individualized speaking tasks within a rigid academic schedule [17]. Zaim et al. [21] observed that inadequate infrastructure and materials prevent the integration of oral performance tasks in underfunded schools. Institutional barriers also exist, including misalignment between policy and practice. Teachers often feel pressured to prioritize summative testing formats that favor accuracy and rote memorization over fluency and spontaneity [7]. The lack of localized and developmentally appropriate speaking tasks further compounds the issue. Ramos and Villena [14] note that most assessment models are designed for secondary learners, leaving elementary teachers without suitable benchmarks or task samples.

Recent studies suggest potential solutions, including sustained professional development, instructional resource support, and policy reform. Firdausi [9] and Alam et al. [2] highlight the importance of assessment literacy programs that help teachers design, adapt, and evaluate authentic tools with greater confidence and consistency. Integrating multiple assessment strategies—including rubrics, peer review, portfolios, and digital tools—can improve reliability and promote fairness, especially when aligned with curriculum standards [11]. Abduh [1] also underscores the value of oral portfolios as part of formative evaluation, supporting learner autonomy and long-term progress tracking.

The reviewed literature affirms that authentic speaking assessment represents a decisive shift toward learner-centered and contextually relevant evaluation. Grounded in theory and enriched by practice, authentic assessment fosters real-world communication, learner reflection, and instructional alignment. However, its full potential can only be realized if



systemic barriers—such as lack of training, insufficient time, and limited resources—are addressed. This study builds on these findings by investigating the actual implementation of authentic assessment tools among Grade 5 English teachers in the Bacuag District, offering data-driven insights that may inform teacher development, resource planning, and curricular alignment in similar educational contexts.

#### **IV. RESULTS AND DISCUSSION**

##### **Profile of the Respondents**

As shown in Table I, the majority of the 19 Grade 5 English teachers in Bacuag District were young (aged 20–30) and early in their careers (1–5 years in service), suggesting openness to innovative strategies like authentic assessment but a potential need for support and training [14]. Most were female (63.16%), reflecting national trends in Philippine elementary education; while gender is not a primary determinant, it may influence classroom interaction styles [19]. Most respondents held bachelor's degrees, with few having pursued graduate studies. Teachers with advanced qualifications tend to demonstrate better alignment with authentic assessment due to stronger theoretical and pedagogical grounding [20]. Teaching experience ranged from novice to veteran, but all groups face systemic challenges such as limited time and resources, which affect consistent implementation [3].

**TABLE I: Profile of the Respondents**

Profile	Categories	Frequency	Percentage
Age	20-30	9	47.37
	31-40	6	31.57
	41-50	2	10.53
	51-60	2	10.53
Sex	Female	12	63.16
	Male	7	36.84
Highest Educational Attainment	Bachelor's Degree	15	78.95
	Master's Degree	4	21.05
Number of Years in Service	less than 1 year	1	5.26
	1-5 years	7	36.84
	6-10 years	3	15.79
	11-15 years and above	8	42.11

##### **Authentic Language Assessment Tools Implemented**

Grade 5 English teachers in the Bacuag District reported using various authentic assessment tools to evaluate speaking skills. These included the Philippine Informal Reading Inventory (Phil-IRI), Test Your Language, oral recitations, role-playing, project-based tasks, rubrics, observation checklists, guided oral explanations, question-and-answer drills, impromptu speaking tasks, group discussions, peer assessments, dialogues, reflection sheets, and video recordings. These tools reflected a combination of standardized, performance-based, and interactive approaches aimed at capturing real-world language use. Teachers used them to assess fluency, comprehension, spontaneity, and collaboration while promoting learner autonomy through reflection and peer feedback. These practices align with studies emphasizing that authentic assessment enhances communicative competence and learner engagement [1, 8].



### Extent of Authentic Language Assessment Tools Implementation in Terms of Characteristics

Table II shows that Grade 5 English teachers in the Bacuag District demonstrated moderate implementation of authentic assessment characteristics ( $M = 3.11$ ,  $SD = 0.92$ ), interpreted as “Evident.” The highest-rated item, “Develops students’ learning mastery” ( $M = 3.47$ ), indicates a strong focus on supporting learners’ skill development. The result aligns with Espinosa and Bernardino [8] and Almerino et al. [3], who emphasized that authentic assessment promotes mastery through real-world application and reflective learning.

Conversely, “Requires an in-depth assessment” received the lowest rating ( $M = 2.89$ ), suggesting limited use of more complex tasks. This finding may stem from contextual constraints such as time and workload, as noted by Sumadsad and Mahinay [16]. Overall, the results highlight the need for further support to enhance the depth and consistency of authentic assessment practices in public elementary schools.

TABLE III: Extent of Authentic Language Assessment Tools Implementation in Terms of Characteristics

Characteristics	Mean	Standard Deviation	Verbal Description	Qualitative Interpretation
1. Focus on the material that is essential, big ideas or the skill-specific skills.	3.25	0.89	Strongly Agree	Highly Evident
2. Requires an in-depth assessment.	2.89	0.86	Agree	Evident
3. Can be implemented easily to do in the classroom or in the school environment.	3.29	0.91	Strongly Agree	Highly Evident
4. Places emphasis on the product quality or performance of the single answer.	2.99	1.01	Agree	Evident
5. Develops students’ learning mastery.	3.47	0.96	Strongly Agree	Highly Evident
6. Communicates to students the criteria required.	3.21	0.88	Agree	Evident
7. Gives the import that criteria have been negotiated by students and teachers.	3.05	0.78	Agree	Evident
8. Tests whether students can demonstrate if they have met the established criteria.	2.99	0.92	Agree	Evident
9. Places essence of the task.	3.32	1.04	Strongly Agree	Highly Evident
<b>Grand Average</b>	<b>3.11</b>	<b>0.92</b>	<b>Agree</b>	<b>Evident</b>

Scale	Range	Verbal Description	Qualitative Interpretation
4	3.25-4.00	Strongly Agree	Highly Evident
3	2.50-3.24	Agree	Evident
2	1.75-2.49	Disagree	Slightly Evident
1	1.00-1.74	Strongly Disagree	Not Evident

### Extent of Authentic Language Assessment Tools Implementation in Terms of Guidelines

Table III shows that Grade 5 English teachers in the Bacuag District demonstrated a high level of implementation of authentic assessment tools in line with recognized guidelines ( $M = 3.28$ ,  $SD = 0.89$ ), interpreted as “Highly Evident.” The highest-rated item used checklists to monitor speaking performance ( $M = 3.64$ ), indicating that teachers frequently rely on observable, formative tools. This result supports Espinosa and Bernardino [8] and Sumadsad and Mahinay [16], who emphasized the value of checklists in tracking progress and ensuring consistency.

The lowest-rated item was rating scales ( $M = 3.00$ ), suggesting limited application of tools for assessing fluency, coherence, and other qualitative aspects of oral output. This finding reflects findings by Almerino et al. [3], who noted





that tools requiring nuanced judgment are often underutilized. Although most teachers aligned assessments with competencies and reported findings clearly, the less frequent use of detailed rubrics and rating scales indicates a need for targeted training to strengthen evaluative depth and accuracy.

**TABLE IIIII: Extent of Authentic Language Assessment Tools Implementation in Terms of Guidelines**

Characteristics	Mean	Standard Deviation	Verbal Description	Qualitative Interpretation
1. The assessment measures the DepEd-prescribed competency standards.	3.57	0.72	Strongly Agree	Highly Evident
2. The assessment tool clearly specifies the basic competency it aims to evaluate.	3.39	0.81	Strongly Agree	Highly Evident
3. The assessment tool includes performance indicators for appraising learner outcomes.	3.29	0.95	Strongly Agree	Highly Evident
4. A rubric is used to assess student performance in speaking tasks.	3.04	1.10	Agree	Evident
5. A checklist is used to monitor learners' speaking performance.	3.64	0.86	Strongly Agree	Highly Evident
6. A detailed checklist is used that describes specific indicators or success criteria.	3.01	0.96	Agree	Evident
7. A rating scale is utilized to evaluate the quality of learners' speaking output.	3.00	0.89	Agree	Evident
8. The teacher consistently records, processes, follows up, and reports assessment findings.	3.32	0.84	Strongly Agree	Highly Evident
<b>Grand Average</b>	<b>3.28</b>	<b>0.89</b>	<b>Strongly Agree</b>	<b>Highly Evident</b>

Scale	Range	Verbal Description	Qualitative Interpretation
4	3.25-4.00	Strongly Agree	Highly Evident
3	2.50-3.24	Agree	Evident
2	1.75-2.49	Disagree	Slightly Evident
1	1.00-1.74	Strongly Disagree	Not Evident

**Test of Significant Difference on the Extent of Authentic Language Assessment Tools Implementation When Grouped According to Profile**

Statistical analysis revealed no significant differences in implementing authentic assessment tools when teachers were grouped by age, sex, or years in service, indicating that these demographic factors did not meaningfully influence assessment practices. Although minor variations in mean scores were noted—particularly among more experienced teachers—these differences were not statistically significant. The result supports previous findings that such demographic variables have minimal impact on instructional behavior and assessment strategies [15, 18].

A significant difference, however, was found in the implementation of authentic assessment characteristics based on educational attainment ( $p = 0.048$ ), with teachers holding graduate degrees reporting higher levels of integration. This result affirms the findings of Zaim et al. [21], who emphasized that teachers with advanced academic training are more likely to apply learner-centered and reflective assessment practices. On the other hand, no significant difference was found in implementing assessment guidelines across educational attainment levels. This finding suggests that



procedural aspects such as aligning tools with prescribed competencies and recording outcomes are uniformly practiced regardless of academic qualification, possibly due to standardized institutional expectations or mandated compliance. It is important to note that one category under years in service was excluded from analysis due to having only one respondent, and two age groups were removed for having only two respondents each. These small sample sizes failed to meet the assumptions required for valid inferential testing, particularly those related to normality and homogeneity of variance. Only age and service groups with sufficient representation were included in the analysis to maintain statistical reliability.

**TABLE IVV: Significant Difference on the Extent of Authentic Language Assessment Tools Implementation When Grouped According to Profile**

Profile	Authentic Language Assessment Tools	p-value	Decision	Interpretation
Age	Characteristics	0.045 <sup>A</sup>	Do not Reject $H_0$	Not Significant
	Guidelines	0.074 <sup>A</sup>	Do not Reject $H_0$	Not Significant
Sex	Characteristics	0.102 <sup>A</sup>	Do not Reject $H_0$	Not Significant
	Guidelines	0.092 <sup>A</sup>	Do not Reject $H_0$	Not Significant
Highest Educational Attainment	Characteristics	0.048 <sup>A</sup>	Reject $H_0$	Significant
	Guidelines	0.092 <sup>A</sup>	Do not Reject $H_0$	Not Significant
Number of Years in Service	Characteristics	0.104 <sup>B</sup>	Do not Reject $H_0$	Not Significant
	Guidelines	0.076 <sup>B</sup>	Do not Reject $H_0$	Not Significant

<sup>A</sup>Tested using Mann-Whitney U Test, <sup>B</sup>Tested using Kruskal-Wallis H Test

## V. CONCLUSION

The demographic profile of the Grade 5 English teachers in the Bacuag District reflected a young and relatively early-career workforce. Most teachers were aged 20–30 and had 1–5 years of teaching experience. Most were female and held bachelor's degrees, with a smaller proportion possessing graduate-level education. This diversity in experience and qualification shaped the general assessment landscape, though statistical analysis showed that most demographic variables did not significantly influence implementation.

Teachers employed various authentic assessment tools for speaking, including oral recitations, role-playing, project-based tasks, group discussions, peer assessments, checklists, rubrics, and reflection-based tools such as video recordings. These tools reflect efforts to integrate both structured and performance-based assessments, promoting real-world application and communicative competence.

The extent of authentic language assessment implementation was rated as Evident in terms of characteristics and Highly Evident in terms of guidelines. Teachers consistently aligned assessments with curriculum standards, monitored learner performance, and supported mastery through practical, learner-centered tasks. However, more complex tools—such as rating scales and detailed rubrics—were less frequently utilized, suggesting the need for further support in their application.

Significant differences in implementation were found only about educational attainment, where teachers with graduate degrees demonstrated greater alignment with the core characteristics of authentic assessment. No significant differences



were observed when grouped by age, sex, or years in service. This finding highlights the role of advanced academic preparation in promoting deeper, reflective assessment practices.

In conclusion, the implementation of authentic language assessment tools in the Bacuag District was evident and procedurally consistent, with strong adherence to guidelines and a clear focus on student learning. To further strengthen speaking assessment, targeted training and support are recommended, particularly in using evaluative tools that require advanced judgment and detailed feedback.

## **VI. ACKNOWLEDGMENT**

The authors sincerely extend their deepest gratitude to all who contributed to the successful completion of this study. Special appreciation is given to their families and loved ones for their unwavering support, patience, and encouragement throughout the most challenging phases of the research journey. Their understanding and presence provided the strength and motivation to continue. The authors also thank the Schools Division Superintendent of Surigao del Norte for granting permission to conduct this study. Gratitude is likewise extended to Surigao del Norte State University, their academic institution, for fostering an environment of critical reflection and intellectual growth that greatly enriched this endeavor. Above all, the authors offer their deepest thanks to God Almighty for granting the wisdom, strength, and perseverance to complete this work. This study is dedicated to all who have stood by the authors in their pursuit of academic and personal growth.

## **REFERENCES**

- [1]. Abduh, A. (2021). EFL teachers' beliefs and practices in authentic speaking assessment: A case from Indonesian secondary schools. *Language Teaching Research Quarterly*, 23, 12–25.
- [2]. Alam, A. C., Abas, T., & Saefullah, H. (2024). Teacher's beliefs and practices of authentic assessment in EFL speaking classes: A case study. *Jurnal Ilmiah Wahana Pendidikan*, 10(7), 183–192.
- [3]. Almerino, P. M., Ocampo, J. M. R., Relloso, S. A., & Busico, M. R. M. (2020). Academic performance and self-efficacy of selected college students. *International Journal of Innovation, Creativity and Change*, 11(8), 310–324.
- [4]. British Council & Social Weather Stations. (2023). English proficiency self-assessment among Filipino adults [Survey report]. As cited in Hernando-Malipot, M. (2025, June 15). 'Spokening' English? Yes, but not quite well. *BusinessMirror*. <https://businessmirror.com.ph/2025/06/15/spokening-english-yes-but-not-quite-well>
- [5]. Capacete, Ma. Pamela. (2020). Case Analysis of the Assessment Practices of Oral Communication Teachers in a Private School in the Philippine Setting. *Modern Journal of Studies in English Language Teaching and Literature*. 1. 42-59. 10.56498/11201990.
- [6]. Chionidou-Moskofoglou, M., Doukakis, S., & Lappa, A. (2021). The use of e-portfolios in teaching and assessment.
- [7]. Dizon, G. (2021). Enhancing speaking skills through performance-based tasks and video assessment in the EFL classroom. *Asian Journal of English Language Studies*, 9(2), 55–74.
- [8]. Espinosa, C. M., & López, G. E. (2020). Assessment considerations for young English language learners. In C. M. Espinosa (Ed.), *Getting it right for young children from diverse backgrounds: Applying research to improve practice* (2nd ed., pp. 143–170). Pearson Education.
- [9]. Firdausi, W. (2024). Teachers' implementation of authentic speaking assessment: A case study of senior high schools in Ambon. *HUELE: Journal of Applied Linguistics, Literature and Culture*, 4(1), 27–39.
- [10]. Gao, L., Blinder, E. B., Barnes, A., Song, K., Clegg, T., Vitak, J., & Chetty, M. (2025). Creating and evaluating privacy and security micro-lessons for elementary school children. *arXiv*.
- [11]. Kim, H., & Jang, E. Y. (2022). Validity and reliability concerns in speaking assessments: A mixed-methods study. *Language Assessment Quarterly*, 19(1), 1–21.
- [12]. Lu, H.-C., Lin, J.-K., Lin, H.-Y., Wang, C.-C., & Chen, B. (2025). Advancing automated speaking assessment leveraging multifaceted relevance and grammar information.





- [13]. Permatasari, N. H., & Nurjati, N. (2023). Authentic assessment on students' speaking skills in project-based learning form of digital storytelling. In InCOLLT 2023 Proceedings.
- [14]. Ramos, A. R., & Villena, M. A. (2023). Collaborative oral assessment practices in the Philippine classroom: Towards student agency in feedback. *Philippine Journal of Language Teaching and Assessment*, 5(1), 12–27.
- [15]. Reynolds, D., & Zupan, B. (2022). Supporting teachers in implementing authentic assessments: The role of collaborative inquiry. *Journal of Teacher Education*, 73(4), 389–402.
- [16]. Sumadsad, L. C., & Mahinay, C. P. (2022). Challenges in implementing performance-based assessment in elementary schools. *Philippine Journal of Education and Learning*, 15(3), 33–46.
- [17]. Sundari, A., & Sabarun, S. (2023). The implementation of authentic assessment in English language teaching. *Jurnal Ilmu Pendidikan Nasional*, 1(2), 83–88.
- [18]. Tran, T. Q., & Nguyen, H. T. T. (2022). Authentic assessment in English language teaching: Practices and challenges in Vietnamese high schools. *REiLA: Journal of Research and Innovation in Language*, 4(1), 13–25.
- [19]. Vercellotti, M. L. (2021). Constructing analytic rubrics for assessing open-ended tasks in the language classroom. *TESL-EJ*, 24(4), Article n4.
- [20]. Wiggins, G. (1998). *Educative assessment: Designing assessments to inform and improve student performance*. Jossey-Bass.
- [21]. Zaim, M., Refnaldi, R., & Arsyad, S. (2020). Authentic assessment for speaking skills: Problem and solution for English secondary school teachers in Indonesia. *International Journal of Instruction*, 13(3), 587–604.

