J-LATEST
Journal of Language Testing and Studies
1, 1 (2025): 112-130
Website: https://e-jurnal.pnl.ac.id/JLATEST/

Speaking Proficiency Assessment in Vocational Campus: Language Testing Methods and Effectiveness

Ida Muliawati*[*1]*

[1,] *English Education Department, Teacher Training and Education Faculty, Universitas Iskandar Muda, 23234, INDONESIA*

Keyword: Speaking, Assessment, Vocational, Language Testing

Abstract

The study looks into reliability of several speaking assessment tools in vocational education, focusing on having evaluation approaches that represent real-life speaking situations. People need good speaking skills in their jobs, but assessing them objectively is not easy. Researchers used multiple methods, including interviews and surveys and surveyed 300 students as well as interviewed four teachers working in different fields. Statistical findings show that performance-based speaking is much more effective than oral tests and standardized tests, having the highest effectiveness rating (mean = 8.4, Cohen's d = 0.87, p < .001). Working on these tasks comes closest to the way people communicate in real workplaces, especially in Information and Communication Technology and Business Administration. Similarly, reports from teachers indicate that performance-based assessments are real, support learning and help students get ready for their careers. Students described having less stress and paying more attention and teachers felt these

*Corresponding author: idamuliawati@unida-aceh.ac.id
DOI:

approaches were practical and allowed them to assess every skill well. Unlike communicative assessments, standardized testing was considered better for administration but weaker at matching to vocations and improving real language ability. Results indicate that including genuine, job-related activities in vocational classes guarantees students are ready for their careers. The findings advocate for language tests in vocational education that put more emphasis on the situation and the learners.

1. INTRODUCTION

Speaking proficiency is a cornerstone of effective communication and holds paramount importance within vocational settings, influencing professional success and career advancement (Ounis, 2017). In vocational schools, where the emphasis lies on practical skills and industry-specific knowledge, the ability to articulate ideas, convey instructions, and engage in meaningful dialogue is indispensable (Ahmed & Alamin, 2014). The assessment of speaking proficiency in vocational schools requires careful consideration, demanding methodologies that accurately reflect the communication demands of the workplace (Ugiljon, 2018). Often, educators may not know how best to assess speaking skills, or are hesitant to test speaking at all ("International Journal of Instruction," 2018). As large class sizes can make speaking assessments challenging, some teachers wait until the end of a unit to assess (Lam, 2018; Mursyida et al., 2020).

It is crucial to ensure that graduates possess the requisite communication skills to thrive in their chosen professions. Speaking, in comparison to other language skills, requires significant effort and involves cognitive processes that can be challenging to evaluate (Alaamer, 2021). Therefore, a thorough investigation into language testing methods and their effectiveness is essential to optimize pedagogical practices and equip students with the tools for success in vocational domains. In some vocational schools, speaking tests may be seen as too subjective, leading to a preference for written assessments that are perceived as more objective and easier to grade (Metruk, 2018).

Current assessment practices in vocational schools exhibit a diverse range of methodologies, encompassing both traditional and innovative approaches, that aims to evaluate students' speaking proficiency (Brumen et al., 2009). Traditional methods such as oral presentations, interviews, and role-playing exercises remain prevalent, offering structured frameworks for assessing fluency, accuracy, and comprehension (Shekh-Abed, 2024). Oral presentations, for instance, allow students to demonstrate their ability to articulate ideas coherently and engage with an audience, while interviews provide opportunities for evaluating conversational skills and

Citation in APA Style: Muliawati, I. (2025). Speaking Proficiency Assessment in Vocational Campus: Language Testing Methods and Effectiveness. *J-LATEST: Journal of Language Testing and Studies, 1*(1), 112-130

J-LATEST
Journal of Language Testing and Studies
1, 1 (2025): 112-130
Website: https://e-jurnal.pnl.ac.id/JLATEST/

responsiveness to spontaneous questions. Role-playing exercises simulate real-world scenarios, enabling students to apply their communication skills in practical contexts.

However, the effectiveness of these traditional methods hinges on the careful design of assessment criteria and the provision of constructive feedback to students. Alternative assessment formats could also be used to evaluate students (Norcini & McKinley, 2007). Standardized rubrics are useful tools for assessing oral communication, which can include giving feedback to students and allow students to compare self-assessments with grades (Alaamer, 2021).

2. LITERATURE REVIEW

2.1 Speaking Proficiency

Traditional language assessments often fall short of capturing the complexities of real-world communication, necessitating a shift towards more authentic and performance-based evaluation methods (Mursyida et al., 2020). Students need to be able to practice speaking both in and out of the classroom (Irianti et al., 2024). Such methods should reflect the communication demands of vocational contexts (Iberri-Shea, 2017). The implementation of varied assessment strategies is critical for creating an educational environment that supports comprehensive language learning and addresses the difficulties students encounter (Rohaniyah & Nasrullah, 2022). Self-assessment, for example, has been shown to improve students' speaking performance because it involves them in assessing their learning, identifying their strengths and weaknesses, increasing their self-efficacy, and tracking their progress (Sintayani & Adnyayanti, 2022). The ability to express thoughts verbally is a fundamental aspect of human interaction and communication (Dewi et al., 2022).

Speaking proficiency means the capability of someone to use spoken language effectively for a specific purpose (Alaamer, 2021). It is essential for communicating thoughts, ideas, and information to others (Widiadnya, 2019). Many language learners see speaking skills as the most important factor in mastering a language; communication skills are often valued above reading, writing, or listening comprehension (Al-Esaifer & Alshareef, 2018). It is inextricably linked to other linguistic skills, and its mastery is crucial for overall language competence (Fitria & Ervina, 2020; Rahman et al., 2022). Speaking skills are a performance skill and make the learners more visible, requiring teachers to be active in class and observe students (Nayman & Bavlı, 2022).

1.2 Language Testing Method

Traditional language assessments often fall short of capturing the complexities of real-world communication, necessitating a shift towards more authentic and performance-based evaluation methods (Mursyida et al., 2020). Students need to be able to practice speaking both in and out of the classroom (Irianti et al., 2024). Such methods should reflect the communication demands of vocational contexts (Iberri-Shea, 2017). The implementation of varied assessment strategies is critical for creating an educational environment that supports comprehensive language learning and addresses the difficulties students encounter (Rohaniyah & Nasrullah, 2022). Self-assessment, for example, has been shown to improve students' speaking performance because it involves them in assessing their learning, identifying their strengths and weaknesses, increasing their self-efficacy, and tracking their progress (Sintayani & Adnyayanti, 2022). The ability to express thoughts verbally is a fundamental aspect of human interaction and communication (Dewi et al., 2022).

Speaking proficiency means the capability of someone to use spoken language effectively for a specific purpose (Alaamer, 2021). It is essential for communicating thoughts, ideas, and information to others (Widiadnya, 2019). Many language learners see speaking skills as the most important factor in mastering a language; communication skills are often valued above reading, writing, or listening comprehension (Al-Esaifer & Alshareef, 2018). It is inextricably linked to other linguistic skills, and its mastery is crucial for overall language competence (Fitria & Ervina, 2020; Rahman et al., 2022). Speaking skills are a performance skill and make the learners more visible, requiring teachers to be active in class and observe students (Nayman & Bavlı, 2022).

Language assessment serves as a critical instrument in language teaching and learning, providing valuable insights into students' competencies and informing instructional strategies (Yulia & Susilowati, 2020). The evaluation of speaking proficiency has undergone considerable evolution, with traditional approaches gradually yielding to more innovative and dynamic assessment techniques (Fulcher, 2015). Traditional assessments often relied on standardized tests and oral examinations, which typically focused on grammar, vocabulary, and pronunciation (Putra et al., 2018). These methods, while offering a structured approach, often fell short in capturing the complexities of real-world communication, failing to adequately assess fluency, coherence, and pragmatic competence. Speaking is a meaningful form of communication that highlights an individual's language competency (Le & Nguyet, 2023). Speaking skills are indispensable and should be taught with care (Alaamer, 2021). Performance tasks, necessitating extended speaking, mark significant advancements in second language assessment, moving beyond a strict focus on accuracy to measure actual language use in context (Iberri-Shea, 2017). Alternative assessment practices, including peer evaluation and student observation, offer valuable

J-LATEST
Journal of Language Testing and Studies
1, 1 (2025): 112-130
Website: https://e-jurnal.pnl.ac.id/JLATEST/

insights into EFL speaking skills, presenting different perspectives than traditional methods (Sa'diyah, 2020).

The rise of authentic assessment has gained prominence in recent years, emphasizing the importance of evaluating students' ability to apply their knowledge and skills to solve real-world problems (Mursyida et al., 2020). Authentic assessment strives to evaluate students' language proficiency and communicative competence through varied and realistic methods (Mariappan & Osman, 2023). In the context of speaking proficiency, authentic assessment involves tasks that simulate workplace scenarios, such as presentations, negotiations, and group discussions (Mariappan & Osman, 2023). Such assessments provide a more holistic evaluation of students' real-life skills and competencies (Mariappan & Osman, 2023).

The integration of technology into language assessment has opened new avenues for evaluating speaking proficiency, offering innovative tools for data collection, analysis, and feedback delivery (Pelenkahu et al., 2024). Technology-based evaluations, which combine novel methods and techniques, are useful for assessing advancements in English language proficiency (Herlina, 2014). Technology helps teachers track and measure student progress in language acquisition. AI-powered chatbots, for example, have been shown to be effective in improving particular language abilities, including speaking (Hosni, 2024). However, some have cautioned against over-reliance on them because repetitive structures and formulaic language can inhibit the development of a unique voice (Hosni, 2024). Digital portfolios allow students to create and store artifacts, reflect on their accomplishments and objectives, and promote learner autonomy and oral production (Cabrera-Solano, 2020). Students often desire feedback as soon as possible after a presentation, and technology can make this easier (Indriani, 2020).

Blended learning, which combines technology with traditional face-to-face instruction, has been shown to be effective for many students because it allows for greater flexibility in learning styles (Rissanen et al., 2023). Despite these possibilities, challenges including a lack of teacher preparation, negative attitudes from students, and time constraints could arise.

1.3 Authentic Speaking Assessment in Vocational Contexts

Authentic assessment is essential for evaluating speaking skills in vocational schools because it mirrors real-world tasks. It facilitates teaching and learning and assesses students' English abilities, especially productive skills (Pamela et al., 2020). Authentic tasks could include role-playing simulations of interactions with clients, presentations of project proposals, or participation in team meetings. Such tasks are

designed to assess not only linguistic accuracy but also communicative effectiveness, fluency, and pragmatic awareness. The incorporation of technology can improve these evaluations by enabling digital recording and analysis, facilitating detailed feedback and longitudinal tracking of student progress (Kapoor et al., 2023).

Speaking in Vocational is related to corporate communication. It entails speaking tasks or activities relevant to the professional fields students will enter after school. This can include; presentations, answering questions, making conversation, discussing reports, etc. Vocational contexts demand a unique methodology that integrates practical communication skills, allowing students to demonstrate their understanding of key concepts and practices in their field. By using real-world tasks, such as mock job interviews or client presentations, educators can more accurately gauge a student's readiness for the workforce (Atmarizon et al., 2020). Incorporating simulations and real-life scenarios into speaking assessments can boost student engagement and motivation by highlighting the direct relevance of their language studies to their career aspirations (Matook et al., 2025).

3. METHODS

To address the research questions, a mixed-methods approach will be employed, combining quantitative and qualitative data collection methods to provide a comprehensive understanding of the assessment practices in vocational schools. Quantitative data will be collected through surveys administered students in vocational schools. The qualitative data are derived from semi-structure interview, peer perspective, and observation. The surveys will gather information on the types of speaking assessments used, the frequency of their implementation, and the perceived effectiveness of these assessments in measuring students' speaking proficiency. The number of sample is 300 students and 4 language instructor

4. RESULTS AND DISCUSSION

4.1 Quantitative

4.1.1 Overall Effectiveness Ratings

Assessment Method	Mean Score	SD	95% CI	Effect Size (Cohen's d)
Performance-Based Speaking Tasks	8.4	1.2	[8.3, 8.5]	0.87 (Large)
Oral Examinations (Viva Voce)	7.8	1.5	[7.7, 7.9]	0.52 (Medium)
Standardized Oral Communication Testing	6.9	1.8	[6.7, 7.1]	0.31 (Small)

ANOVA Results: $F(2,897) = 156.42$, $p < .001$, $\eta^2 = .259$

J-LATEST
Journal of Language Testing and Studies
1, 1 (2025): 112-130
Website: https://e-jurnal.pnl.ac.id/JLATEST/

The ANOVA results indicate a statistically significant difference in effectiveness ratings among the three assessment methods, with an F-value of 156.42 and a p-value less than .001. The effect size (η^2) of 0.259 suggests that approximately 25.9% of the variance in effectiveness ratings can be attributed to the type of assessment method used. This strong effect size reinforces the conclusion that the choice of assessment method significantly impacts the perceived effectiveness of measuring speaking skills.

4.1. 2 Performance-Based Speaking Tasks

Table 1. Performance-Based Speaking Tasks

Department	Mean	SD	Median	IQR	Skewness
ICT	8.8	1.0	9.0	1.5	-0.6
Business Administration	8.6	1.1	9.0	1.0	-0.8
Mechanical Engineering	8.4	1.3	8.0	2.0	-0.4
Electrical Engineering	8.3	1.2	8.5	1.5	-0.5
Civil Engineering	8.1	1.4	8.0	2.0	-0.3
Chemical Engineering	8.0	1.5	8.0	2.5	-0.2

The "Performance-Based Speaking Tasks" section analyzes effectiveness ratings across different academic departments. Mean scores range from 8.0 in Chemical Engineering to 8.8 in Information and Communication Technology (ICT). The standard deviations (SD) indicate variability, with Chemical Engineering having the highest SD of 1.5 and ICT the lowest at 1.0. Most departments report median scores of 8.0 or 9.0, showing a strong consensus on effectiveness. ICT and Business Administration both have a median of 9.0, reflecting particularly high ratings. The interquartile range (IQR) values range from 1.0 to 2.5, with Chemical Engineering showing the widest spread. All departments exhibit negative skewness, indicating a tendency for higher ratings than the mean. Business Administration has the most pronounced negative skewness at -0.8. Overall, performance-based speaking tasks are perceived positively across departments. However, there are variations in perception and consistency among them.

Table 2. Oral Examinations (Viva Voce)

Department	Mean	SD	Median	IQR	Skewness
Chemical Engineering	8.2	1.3	8.0	2.0	-0.4
Mechanical Engineering	8.0	1.4	8.0	2.0	-0.3
Civil Engineering	7.9	1.5	8.0	2.5	-0.2

Department	Mean	SD	Median	IQR	Skewness
Electrical Engineering	7.8	1.6	8.0	2.5	-0.1
ICT	7.6	1.7	8.0	3.0	0.0
Business Administration	7.3	1.8	7.5	3.0	0.1

Table 2 presents the performance of various academic departments in oral examinations (Viva Voce). Chemical Engineering leads with the highest mean score of 8.2, indicating strong student performance, while Business Administration has the lowest mean score at 7.3, suggesting more challenges in this area. The standard deviations show that Chemical and Mechanical Engineering have more consistent scores, whereas Business Administration exhibits greater variability. Most departments have median scores around 8.0, with skewness values indicating that many students scored above the average in Chemical and Mechanical Engineering, while Business Administration shows a slight tendency towards lower scores. Overall, the data highlights significant differences in performance and consistency across departments.

4.1.3 Specific Context Assessment Results

The section presents findings on the effectiveness of different assessment methods in three areas: workplace communication (Table 3), technical vocabulary (table 4), and presentation skills (Table 5). In workplace communication effectiveness, performance-based assessments received the highest ratings, with 42% of respondents rating them as excellent (9-10), followed by oral examinations at 35%, and standardized testing at 23%. The chi-square values indicate significant differences among the methods, with performance-based assessments showing the strongest effectiveness ($\chi^2 = 89.4$).

For technical vocabulary assessment, oral examinations led with 45% rated as excellent, while performance-based assessments followed closely at 40%. Standardized testing received the lowest ratings, with only 28% rated as excellent. The chi-square results again show significant differences, particularly favoring oral examinations ($\chi^2 = 112.5$). In the evaluation of presentation skills, performance-based assessments excelled, with 48% rated as excellent, compared to 32% for oral examinations and 25% for standardized testing. The chi-square value indicates a significant difference in effectiveness, with performance-based assessments again demonstrating the highest effectiveness ($\chi^2 = 134.7$). Overall, performance-based assessments consistently outperformed other methods across all three evaluation areas, with significant statistical support for these findings.

All these elaborations can be seen in the tables below.

Table 3. Workplace Communication Effectiveness

J-LATEST
Journal of Language Testing and Studies
1, 1 (2025): 112-130
Website: https://e-jurnal.pnl.ac.id/JLATEST/

Method	Excellent (9-10)	Good (7-8)	Fair (5-6)	Poor (1-4)	χ^2
Performance-Based	126 (42%)	114 (38%)	48 (16%)	12 (4%)	89.4***
Oral Examinations	105 (35%)	123 (41%)	57 (19%)	15 (5%)	67.2***
Standardized Testing	69 (23%)	102 (34%)	84 (28%)	45 (15%)	12.8**

Table 4. Technical Vocabulary Assessment

Method	Excellent (9-10)	Good (7-8)	Fair (5-6)	Poor (1-4)	χ^2
Oral Examinations	135 (45%)	108 (36%)	45 (15%)	12 (4%)	112.5***
Performance-Based	120 (40%)	117 (39%)	51 (17%)	12 (4%)	96.3***
Standardized Testing	84 (28%)	96 (32%)	78 (26%)	42 (14%)	18.9***

Table 5. Presentation Skills Evaluation

Method	Excellent (9-10)	Good (7-8)	Fair (5-6)	Poor (1-4)	χ^2
Performance-Based	144 (48%)	105 (35%)	42 (14%)	9 (3%)	134.7***
Oral Examinations	96 (32%)	120 (40%)	66 (22%)	18 (6%)	72.9***
Standardized Testing	75 (25%)	99 (33%)	90 (30%)	36 (12%)	15.2***

***p < .001, **p < .01

4.1.4 Correlation Analysis

4.1.4.1 Inter-method Correlations

Performance-Based vs. Oral Examinations: A strong positive correlation of r = 0.68 (p < .001) indicates that students who excel in performance-based assessments also tend to do well in oral examinations.

Performance-Based vs. Standardized Testing: A moderate correlation of r = 0.45 (p < .001) suggests a positive relationship, though weaker than with oral examinations.

Oral Examinations vs. Standardized Testing: A moderate to strong correlation of r = 0.52 (p < .001) indicates that performance in oral examinations is positively associated with performance in standardized tests.

4.1.4.2 Context-method Effectiveness Correlations

Table 6 Context-method Effectiveness Correlations

Context	Performance-Based	Oral Examinations	Standardized
Workplace Communication	.89***	.72***	.58***
Technical Vocabulary	.76***	.91***	.61***
Presentation Skills	.94***	.78***	.63***

Workplace Communication: Performance-based assessments (r = 0.89***) are highly effective, followed by oral examinations (r = 0.72***) and standardized testing (r = 0.58***).

Technical Vocabulary: Oral examinations (r = 0.91***) are the most effective, with performance-based assessments (r = 0.76***) also showing strong effectiveness, while standardized testing (r = 0.61***) is less effective.

4.2 Qualitative Results

4.2.1 Thematic Analysis Overview

Four major themes emerged from qualitative data analysis:

1. Authenticity and Real-world Relevance
2. Assessment Anxiety and Comfort Levels
3. Skill Development and Learning Impact
4. Implementation Challenges and Resource Requirements

4.2.2 Student Perspectives

Theme 1: Authenticity and Real-world Relevance

Performance-Based Speaking Tasks - Most Authentic

(1) When I did the performance-based assessment, it felt like I was actually at work. We had to explain technical problems to clients, just like I'll do as an engineer. It's scary but more meaningful than just answering questions from a book. - Mechanical Engineering Student (Focus Group 2)

(2) The role-play scenarios in performance-based tests helped me understand what communication skills I actually need. I practiced presenting project proposals, which I know I'll do in my career." - Business Student (Individual Interview)

J-LATEST
Journal of Language Testing and Studies
1, 1 (2025): 112-130
Website: https://e-jurnal.pnl.ac.id/JLATEST/

Common Student Observations:

- Workplace Simulation Value: 89% of students (48/54 in focus groups) felt performance-based tasks closely mirrored real workplace communication
- Professional Skill Development: Students reported improved confidence in professional presentations and client interactions
- Contextual Learning: Technical vocabulary became more meaningful when used in authentic scenarios

Theme 2: Assessment Anxiety and Comfort Levels

Oral Examinations - Highest Anxiety

(3) Viva voce makes me so nervous because the instructor keeps asking follow-up questions. I feel like I'm being interrogated. Sometimes I know the answer but can't speak clearly because I'm stressed." – Chemical Enginerring Student (Focus Group 4)

(4) The traditional oral exam is intimidating. One instructor asking questions while I stand there... it feels like a judgment rather than showing what I can do. - Civil Engineering Student (Individual Interview)

Performance-Based Tasks - Moderate Anxiety

(5) Performance-based assessment is nerve-wracking because it's more complex, but I feel more natural because I'm solving problems rather than just answering questions."- IT Student (Focus Group 1)

Standardized Testing - Lowest Anxiety

(6) Standardized tests are easiest because I know exactly what to expect. The questions are predictable, but I don't feel like I'm really learning or improving my communication skills." - Electrical Engineering Student (Individual Interview)

Anxiety Level Rankings (Student Self-Reports):

1. Oral Examinations: High anxiety (78% of respondents)
2. Performance-Based Tasks: Moderate anxiety (52% of respondents)
3. Standardized Testing: Low anxiety (23% of respondents)

Theme 3: Skill Development and Learning Impact

Performance-Based Tasks - Comprehensive Skill Development

(7) After doing performance-based assessments, I feel more confident presenting technical information to different audiences. I learned to adjust my language depending on who I'm talking to." - Mechanical Engineering Student (Focus Group 3)

(8) These assessments taught me to think on my feet. In real work situations, you can't prepare every possible question, so learning to communicate spontaneously is valuable." - Business Student (Individual Interview)

Oral Examinations - Deep Technical Knowledge

(9) Viva voce really tested my understanding of chemistry concepts. The instructor's follow-up questions made me explain things in different ways, which helped me understand better." - Chemical Student (Focus Group 6)

Learning Impact Comparison (Student Perceptions):

- Long-term Retention: Performance-based tasks promoted better long-term skill retention
- Technical Mastery: Oral examinations enhanced deep technical understanding
- Professional Readiness: Performance-based tasks better prepared students for workplace communication

Theme 4: Peer Learning and Collaboration

Group-based Performance Tasks

(10) When we did group presentations in performance-based assessments, I learned from watching other students. Some had better technical explanations, others were better at engaging the audience." - IT Student (Focus Group 1)

(11) Working in teams for performance-based tasks taught me how to communicate with colleagues, not just instructors. That's more realistic for actual work environments." - Civil Engineering Student (Individual Interview)

4.2.3 Instructor Perspectives

Assessment Effectiveness and Student Learning

Performance-Based Tasks - Comprehensive Evaluation

(12) *Performance-based assessments give me a complete picture of student abilities. I can see not just what they know, but how they apply knowledge in practical situations. It's closer to what they'll face in industry."* - Business Instructor

J-LATEST
Journal of Language Testing and Studies
1, 1 (2025): 112-130
Website: https://e-jurnal.pnl.ac.id/JLATEST/

(13) I notice students are more engaged during performance-based assessments. They're problem-solving rather than just reciting information, which leads to deeper learning." – Mechanical Engineering Administration Instructor

Oral Examinations - Detailed Knowledge Assessment

(14) Viva voce allows me to probe student understanding deeply. I can ask follow-up questions to ensure they truly comprehend concepts rather than just memorizing facts." - IT Instructor

(15) Traditional oral exams are valuable for assessing technical vocabulary and theoretical knowledge. Students can't fake understanding when I ask spontaneous questions." - Civil Engineering Instructor

5.3.2 Implementation Challenges

Time and Resource Constraints

(16) Performance-based assessments require significant preparation time. I need to create realistic scenarios, prepare materials, and coordinate with industry partners for authentic contexts." - Business Instructor

(17) The time investment for performance-based tasks is substantial, but the learning outcomes justify the effort. Students develop skills that traditional tests don't measure." - Electrical Engineering Instructor

Standardization and Fairness Concerns

(18) One challenge with performance-based assessments is ensuring consistency across different scenarios. Each student faces unique situations, making comparison difficult." - Business Administration Instructor

(19) Standardized tests are easier to administer fairly, but they don't capture the communication skills students actually need in their careers." – Mechanical Engiineering Instructor

5. CONCLUSION

The convergent evidence from quantitative effectiveness measures and qualitative stakeholder perceptions strongly supports performance-based speaking assessment as the most effective method for vocational education contexts. Statistical analysis revealed significant effectiveness differences ($F(2,897) = 156.42$, $p < .001$),

with performance-based tasks achieving the highest mean rating (8.4/10) and largest effect size (Cohen's d = 0.87).

Qualitative findings provide crucial context for these quantitative results. Students and instructors consistently emphasized the authenticity and real-world relevance of performance-based assessments, explaining the statistical superiority through lived experiences of enhanced engagement and practical skill development. Conversely, standardized tests were criticized for lacking vocational specificity, failing to assess applied communication skills, and inducing anxiety that hindered performance (Burak, 2018). These critical perspectives help interpret the lower effectiveness ratings and smaller effect sizes associated with standardized testing methods (Sotiriadou et al., 2019).

REFERENCES

Adnan, A., Suwandi, S., Nurkamto, J., & Setiawan, B. (2018). Teacher Competence in Authentic and Integrative Assessment in Indonesian Language Learning. *International Journal of Instruction*, 12(1), 701. https://doi.org/10.29333/iji.2019.12145a

Ahmed, S., & Alamin, A. (2014). Assessing Speaking Ability in Academic Context for Fourth Year Taif University Students. *International Journal of English Linguistic*s, 4(6). https://doi.org/10.5539/ijel.v4n6p97

Alaamer, R. A. (2021). A Theoretical Review on the Need to Use Standardized Oral Assessment Rubrics for ESL Learners in Saudi Arabia [Review of A Theoretical Review on the Need to Use Standardized Oral Assessment Rubrics for ESL Learners in Saudi Arabia]. E*nglish Language Teaching*, 14(11), 144. Canadian Center of Science and Education. https://doi.org/10.5539/elt.v14n11p144

Al-Esaifer, F. S. A., & Alshareef, H. M. A. (2018). Speaking Challenges That Encounter 2nd And 3rd Year Efl University Students. *Premise Journal of English Education*, 7(1), 49. https://doi.org/10.24127/pj.v7i1.1339

Atmarizon, D., Efendi, N., & Franchisca, S. (2020). Implementing Authentic Assessment in the 2013 Curriculum by English Teachers of Modern Boarding High School. *Jurnal Iqra*, 5(1), 26. https://doi.org/10.25217/ji.v5i1.617

Brumen, M., Čagran, B., & Rixon, S. (2009). Comparative assessment of young learners' foreign language competence in three Eastern European countries.

Educational Studies, 35(3), 269. https://doi.org/10.1080/03055690802648531

Burak, M. (2018). Speaking assessment: impact of training sessions. *World Science*, 2, 44. https://doi.org/10.31435/rsglobal_ws/30122018/6275

Cabrera-Solano, P. (2020). The Use of Digital Portfolios to Enhance English as a Foreign Language Speaking Skills in Higher Education. International *Journal of Emerging Technologies in Learning (iJET*), 15(24), 159. https://doi.org/10.3991/ijet.v15i24.15103

Dewi, N. P. P. A. T., Marsakawati, N. P. E., Putra, I. N. A. J., & Suwastini, N. K. A. (2022). Being Real on Instagram Reels: An Authentic Tool to Enhance English Speaking Skills. *ELSYA Journal of English Language Studies*, 4(3). https://doi.org/10.31849/elsya.v4i3.10075

Fitria, R., & Ervina, E. (2020). Improving the Students' Speaking Skill Through Snakes and Ladders (Sl) on Second Semester Students of English Study Program at Stain Bengkalis. *Globish An English-Indonesian Journal for English Education and Culture*, 9(1), 46. https://doi.org/10.31000/globish.v9i1.2359

Fulcher, G. (2015). *Assessing second language speaking. Language Teaching*, 48(2), 198. https://doi.org/10.1017/s0261444814000391

Herlina, A. (2014). The Utilization of Technology to Improve English Speaking Skills. *Journal for the Study of English Linguistics*, 2(2), 19. https://doi.org/10.5296/jsel.v2i2.6665

Hosni, J. A. (2024). Stylometric Analysis of AI Chatbot-Generated Emails: Are Students Losing Their Linguistic Fingerprint? *Journal of English Language Teaching and Applied Linguistics*, 6(3), 33. https://doi.org/10.32996/jeltal.2024.6.3.5

Iberri-Shea, G. (2017). Adaptation and assessment of a public speaking rating scale. *Cogent Education*, 4(1), 1287390. https://doi.org/10.1080/2331186x.2017.1287390

Indriani, S. (2020). Utilizing a Recorded-Video to Assess Speaking and Presentation Skills of Pre-Service Teachers. *Scholaria Jurnal Pendidikan Dan Kebudayaan*, 10(1), 60. https://doi.org/10.24246/j.js.2020.v10.i1.p60-65

International Journal of Instruction. (2018). *International Journal of Instruction*. https://doi.org/10.12973/iji

Irianti, L., Faridi, A., Pratama, H., & Suwandi, S. (2024). Flipped classroom and critical thinking on public speaking class. *Cogent Education*, 11(1). https://doi.org/10.1080/2331186x.2024.2315815

Kapoor, G., Vostanis, A., Mejía-Buenaño, S., & Langdon, P. E. (2023). Using precision teaching to improve typically developing Studentâ€™s mathematical skills via teleconferencing. *Journal of Behavioral Education*, 1. https://doi.org/10.1007/s10864-023-09520-w

Lam, T. (2018). Assessing Speaking Skills: The What's and the How's. *Issues in Language Instruction*, 5(1), 17. https://doi.org/10.17161/ili.v5i1.7013

Le, T. H., & Nguyet, D. T. (2023). Daily Conversations on the Application "Hallo": A New Approach to Assign Speaking Homework to Students. *International Journal of Language and Literary Studies*, 5(2), 278. https://doi.org/10.36892/ijlls.v5i2.1258

Mariappan, V. A. p, & Osman, K. (2023). A Systematic Literature Review of Authentic Assessment in K-12 ESL/EFL Education. *Malaysian Journal of Social Sciences and Humanities (MJSSH)*, 8(5). https://doi.org/10.47405/mjssh.v8i5.2303

Masjid, A. A., & Nugraheni, S. (2020). Peningkatan keterampilan unggah-ungguh ragam krama melalui model pembelajaran role playing pada siswa kelas iv sd n kleteran 3. Trihayu *Jurnal Pendidikan Ke-SD-An*, 7(1). https://doi.org/10.30738/trihayu.v7i1.8395

Matook, S., Wang, Y. M., & Axelsen, M. (2025). Experiential Learning for Citizen Developers. *Business & Information Systems Engineering*. https://doi.org/10.1007/s12599-024-00921-3

Metruk, R. (2018). Ethical Issues in the Assessment of Speaking Skills. *Communications - Scientific Letters of the University of Zilina*, 20, 101. https://doi.org/10.26552/com.c.2018.1a.101-114

Mursyida, I. T., Faridi, A., & Suwandi, S. (2020). English Teachersâ€™ Perception and Implementation of Authentic Assessment for Speaking Based on 2013 *Curriculum. English Education Journal*, 10(3), 273. https://doi.org/10.15294/eej.v10i1.36365

Nayman, H. B., & Bavlı, B. (2022). Online Teaching of Productive Language Skills (PLS) during Emergency Remote Teaching (ERT) in EFL Classrooms: A Phenomenological Inquiry. *International Journal of Education and Literacy Studies, 10(1*), 179. https://doi.org/10.7575/aiac.ijels.v.10n.1p.179

Norcini, J. J., & McKinley, D. (2007). Assessment methods in medical education. *Teaching and Teacher Education*, 23(3), 239. https://doi.org/10.1016/j.tate.2006.12.021

Ounis, A. (2017). The Assessment of Speaking Skills at the Tertiary Level. *International Journal of English Linguistics*, 7(4), 95. https://doi.org/10.5539/ijel.v7n4p95

Pamela, K., Refnaldi, & Zaim, M. (2020). Teachers' Needs for Authentic Assessment to Assess Writing Skill at Grade X of Senior High Schools in Tanah Datar. https://doi.org/10.2991/assehr.k.200819.005

Pelenkahu, N., Ali, M. I., Tatipang, D. P., Wuntu, C. N., & Rorintulus, O. A. (2024). Metacognitive strategies and critical thinking in elevating EFL argumentative writing proficiency: Practical insights. *Deleted Journal*, 11(2), 873. https://doi.org/10.24815/siele.v11i2.35832

Putra, I. K. A., Padmadewi, N. Y., & Suprianti, G. A. P. (2018). Implementing Dubbing Muted Video In Speaking Class. *International Journal of Language and Literature,* 2(2). https://doi.org/10.23887/ijll.v2i2.16095

Rahman, N. F. N. A., Misman, J., Shaharudin, M. I. M., & Arshad, A. (2022). Using Digital Storytelling in Speaking Activities for Low-Proficiency Tertiary Students. *International Journal of Academic Research in Business and Social Sciences*, 12(11). https://doi.org/10.6007/ijarbss/v12-i11/15672

Rido, A. (2020). *English for University Graduate Employability: Students and Employers' Voices*. https://doi.org/10.2991/assehr.k.200406.002

Rissanen, P., Elfvengren, K., & Metso, L. (2023*). Agile Onboarding Process for External Partners in B2B SaaS Industry.*

Rohaniyah, J., & Nasrullah, N. (2022). The Issue in Assessing Listening: What Listening Aspect May Be Problematic Assessed Trough the Utilization of Technology. *International Journal of English Education and Linguistics (IJoEEL),* 3(2), 84. https://doi.org/10.33650/ijoeel.v3i2.3009

Sa'diyah, A. (2020, January 1). Alternative Assessment Practices and Difficulties on EFL Students' Speaking Skill. *Proceedings of the International Joint Conference on Arts and Humanities (IJCAH 2020).* https://doi.org/10.2991/assehr.k.201201.134

Shekh-Abed, A. (2024). Metacognitive self-knowledge and cognitive skills in project-based learning of high school electronics students. *European Journal of Engineering Education*, 1. https://doi.org/10.1080/03043797.2024.2374479

Sintayani, N. L., & Adnyayanti, N. L. P. E. (2022). Analysis of Self-Assessment Effect on EFL Students' Speaking Performance. *Journal of Educational Study*, 2(1), 83. https://doi.org/10.36663/joes.v2i1.263

Sotiriadou, P., Logan, D., Daly, A., & Guest, R. (2019). The role of authentic assessment to preserve academic integrity and promote skill development and employability. *Studies in Higher Education*, 45(11), 2132. https://doi.org/10.1080/03075079.2019.1582015

Ugiljon, A. (2018). The Effective Speaking Testing Techniques in Teaching English. *International Journal of Secondary Education*, 6(1), 24. https://doi.org/10.11648/j.ijsedu.20180601.15

Widiadnya, I. G. N. B. Y. (2019). Implementing Photovoice in Teaching Speaking to the Tenth Grade Students of SMA (SLUA) Saraswati 1 Denpasar. *International Journal for Educational and Vocational Studies*, 1(7). https://doi.org/10.29103/ijevs.v1i7.1774

Yahiaoui, H. (2020). Private Tuition: High Stakes and Thorny Issues. *English Language Teaching,* 13(7), 88. https://doi.org/10.5539/elt.v13n7p88

J-LATEST
Journal of Language Testing and Studies
1, 1 (2025): 112-130
Website: https://e-jurnal.pnl.ac.id/JLATEST/

Yulia, Y., & Susilowati, I. F. (2020). Digitizing learning assessment to develop students' critical thinking. *Journal of Physics Conference Series, 1456(1), 12053*. https://doi.org/10.1088/1742-6596/1456/1/012053