

THE IMPLEMENTATION OF PERFORMANCE-BASED ASSESSMENT ON HIGH SCHOOL STUDENTS' SPEAKING ABILITY IN AN EXTRACURRICULAR PROGRAM

Muhammad Bimo Aldinsyah^{*1}, Ratna Andhika Mahaputri², Mahmud Mahmud³
bimoaldinsyah84@gmail.com^{*1}, ratna.andhika@ugj.ac.id²,
mahmood.cliff@gmail.com³

Faculty of Education and Sciences ^{*1, 2, 3}
Swadaya Gunung Jati University^{*1, 2, 3}

Received: July 22, 2026

Accepted: September 18, 2026

Published: September 25, 2026

ABSTRACT

Speaking proficiency remains a critical yet often neglected skill in high school EFL contexts in Indonesia, where conventional assessment practices continue to prioritize written tests over authentic oral performance. While Performance-Based Assessment (PBA) has been proven effective in tertiary education, its implementation in high school English Clubs remains underexplored. This study aimed to examine changes in high school students' speaking ability following PBA, describe students' perceptions of PBA, and explore students' experiences in a six-month program within a compulsory extracurricular English Club at a boarding school. An explanatory sequential mixed-methods design was employed with 40 high school students. Quantitative data were collected through speaking pretests and posttests using an analytic rubric from Brown (2004) and an eight-item questionnaire on a four-point Likert scale, analyzed descriptively. Five students were subsequently interviewed and data were analyzed using thematic analysis following Braun and Clarke (2006). Results indicate a descriptive increase in speaking scores across all five dimensions. Students reported generally positive perceptions, particularly regarding tutor feedback, demonstrating the viability of PBA as a structured assessment approach in extracurricular settings. However, gains in confidence and independent practice varied across proficiency levels, suggesting that PBA may require additional differentiated support in mixed-proficiency compulsory settings.

Keywords: *English Club, Extracurricular Program, High School Students, Performance-Based Assessment, Speaking Ability.*

DOI: 10.31943/wej.v10i2.751

INTRODUCTION

Speaking is widely recognized as one of the most critical yet challenging skills for high school students to master in English language learning (De Jong,

THE IMPLEMENTATION OF PERFORMANCE-BASED ASSESSMENT
ON HIGH SCHOOL STUDENTS' SPEAKING ABILITY
IN AN EXTRACURRICULAR PROGRAM

2023). Proficiency in spoken English has become a practical necessity for academic success, future careers, and social integration. However, despite its acknowledged importance, speaking remains the most neglected skill in classroom assessment, with research confirming that oral competence is often underevaluated in high school education (Bashori et al., 2024). In the Indonesian educational context, conventional assessment practices continue to favor written tests over oral communicative ability, a trend observed in both high schools (Fadila & Trisno, 2025) and higher education (Menggo & Gunas, 2022). When students are not required to demonstrate their speaking ability in a structured and authentic manner, their oral competence risks remaining both underdeveloped and unevaluated (Zaim et al., 2020).

Speaking is also one of the most cognitively demanding skills for high school learners to develop, requiring the simultaneous management of phonological encoding, lexical retrieval, grammatical formulation, discourse organization, and real-time responsiveness to an interlocutor (De Jong, 2023). For learners in Indonesian high schools, this challenge is compounded by limited opportunities for authentic oral practice in regular classrooms (Menggo & Gunas, 2022). Drawing on the model developed by Bachman and Palmer (1996), De Jong (2023) identifies two core components of speaking proficiency: linguistic competence, covering the structural, pragmatic, and sociolinguistic dimensions of spoken language, and strategic competence, involving the communication strategies learners use to handle breakdowns and sustain interaction. These two components work together in any communicative task, from everyday conversation to formal presentations.

In classroom pedagogy, the principles of Communicative Language Teaching (CLT) have shaped the way speaking is taught (Richards, 2006). CLT places communicative competence at the center of instruction rather than focusing narrowly on grammatical accuracy, expecting learners to develop both accuracy and fluency through meaningful interaction and authentic tasks (Nunan, 1991; Richards, 2006). Research consistently demonstrates that despite years of formal English instruction, many Indonesian high school students continue to struggle with effective oral communication, particularly in fluency, pronunciation accuracy, and confidence (Menggo & Gunas, 2022). Nggawu and Thao (2023) found that CLT-based speaking instruction produced measurable improvements in students' oral performance across different learner personality types. These findings reinforce the need for speaking to be assessed through methods genuinely aligned with its communicative nature.

This points to the need for an assessment approach that genuinely reflects the performative nature of speaking. Language assessment scholars argue that speaking should be evaluated through real communicative performance rather than indirect measures of linguistic knowledge, and that a valid speaking assessment must capture what learners can actually do with the language (De Jong, 2023). A central concern in speaking assessment is validity, as task design, the role of the interlocutor, scoring criteria, and rater consistency may all introduce construct-irrelevant variance (Bachman & Palmer, 1996; De Jong, 2023). To reduce these

sources of bias and ensure transparency, the use of a structured analytic rubric is considered essential, providing clear descriptors for each dimension of oral performance and enabling reliable, pedagogically meaningful judgments (Dimova, 2022; Nadolski et al., 2021; Stanley, 2021b). A well-designed rubric specifies dimensions typically including pronunciation, grammar, vocabulary, fluency, and comprehension, with level-specific descriptors that guide raters toward consistent and defensible evaluations (Nadolski et al., 2021; Stanley, 2021b).

Analytic rubrics, which assign independent scores to each dimension, are especially useful in speaking classrooms because they offer students specific, actionable feedback on each area requiring improvement (Brookhart, 2018; Nadolski et al., 2021). Nadolski and colleagues (2021) found that rubric-based formative assessment outperformed conventional rating-scale approaches in improving high school students' oral presentation skills, and that rater training significantly enhanced inter-rater reliability. This view aligns with the broader framework of authentic assessment, which insists that evaluation tasks be grounded in real communicative performance and mirror meaningful language use beyond the classroom (Stanley, 2021a). It is from this foundation that Performance-Based Assessment emerges as a directly applicable approach to the assessment of speaking skills (Dimova, 2022; O Malley & Valdez Pierce, 1996).

Within this context, Performance-Based Assessment (PBA) emerges as a particularly relevant approach for teaching and evaluating speaking skills (Stanley, 2021a). PBA requires learners to actively demonstrate a skill, placing emphasis on both the learning process and the final outcome. In speaking contexts, PBA tasks replicate authentic communicative situations, offering students opportunities to use language purposefully and be evaluated on their actual oral performance (Menggo & Gunas, 2022; Stanley, 2021a). In the Indonesian context, Lestari and Azizah (2021) found that PBA activities enabled learners to apply grammar features in realistic contexts, while Aprianto and Sutarman (2025) found that rubric-based speaking assessment provided a more comprehensive measure of students' oral competence. Empirical evidence consistently indicates that PBA is associated with improvements across key speaking dimensions while also supporting affective outcomes such as self-confidence, learning autonomy, and reduced speaking anxiety (Fadila & Trisno, 2025; Menggo & Gunas, 2022; Plengkham & Wasanasomsithi, 2023). Recent studies in Indonesian high schools have similarly found that PBA supports improvements in speaking performance while enhancing motivation and confidence (Afifah & Widyantoro, 2025).

Extracurricular English Clubs have gained recognition as valuable supplementary learning environments in Indonesian high schools, particularly where classroom time for oral communication is limited (Maros et al., 2023; Octaberlina & Muslimin, 2022). Simbolon, Tampubolon, and Pasaribu (2023) found that English Club participation significantly improved students' speaking ability, with mean scores increasing by 56%, while student engagement rose from 47% to 77% across sessions. Pratiwi (2024) similarly found that 81% of high school students reported increased confidence in speaking following English Club participation, though she also identified persistent challenges including pronunciation difficulties and limited vocabulary retention, issues that a structured,

THE IMPLEMENTATION OF PERFORMANCE-BASED ASSESSMENT
ON HIGH SCHOOL STUDENTS' SPEAKING ABILITY
IN AN EXTRACURRICULAR PROGRAM

rubric-based assessment framework could help address. Octaberlina and Muslimin (2022) found that speaking activities such as speech, storytelling, and group discussion contributed positively to students' oral development in vocabulary, grammar, and pronunciation, while Maros et al. (2023) noted that regular, structured oral practice was especially effective in developing fluency and vocabulary. In boarding school contexts, English Clubs often operate as compulsory programs integrated into the weekly schedule, ensuring participation across all proficiency levels and eliminating the self-selection bias commonly found in voluntary extracurricular studies (Gisma et al., 2025; Salsabila et al., 2025).

Despite these encouraging findings, significant gaps remain in the literature. Most PBA studies have been conducted with university students in formal classroom settings (Menggo & Gunas, 2022; Plengkham & Wasanasomsithi, 2023), while research at the high school level is limited and focuses primarily on regular classrooms rather than extracurricular programs (Fadila & Trisno, 2025; Putri et al., 2024). Studies in English Club settings have rarely incorporated a systematic, rubric-based assessment framework, and the unique context of boarding schools where English Clubs are compulsory has received little attention in the literature. Consequently, how PBA can be implemented to assess and support students' speaking ability within a compulsory high school English Club at a boarding school remains largely unexplored.

The present study seeks to fill this gap by examining the implementation of PBA through a speaking performance activity conducted within a compulsory high school English Club at a boarding school. Unlike voluntary English Clubs in public schools, this setting includes students across all proficiency levels, eliminating self-selection bias. Students prepared and delivered spoken English across a range of communicative tasks over six months, culminating in a performance day that served as both an authentic language event and a summative speaking assessment, evaluated using an analytic rubric adapted from Brown (2004).

This study employs an explanatory sequential mixed-methods design (Creswell & Plano Clark, 2017), combining quantitative analysis of speaking scores and questionnaire data with qualitative interviews analyzed using thematic analysis (Braun & Clarke, 2006). It is guided by three research questions: how students' English speaking scores change following the implementation of Performance-Based Assessment in the English Club extracurricular program; how students perceive the implementation of PBA in the English Club extracurricular program; and how students describe their experiences and reflections after participating in the six-month PBA program. Correspondingly, the study aims to (1) describe changes in high school students' English speaking scores following PBA implementation, (2) describe students' perceptions of PBA as a speaking assessment approach, and (3) explore students' experiences and reflections on participating in the program and its perceived effects on their speaking development.

RESEARCH METHOD

Research Design

This study employs a mixed methods research design. A mixed methods approach is a procedure for collecting, analyzing, and integrating both quantitative and qualitative data in a single study to better understand a research problem than either type of data alone (Creswell & Plano Clark, 2017). Specifically, this study uses an explanatory sequential design, in which quantitative data are collected and analyzed first, followed by qualitative data to help explain the quantitative results (Creswell & Guetterman, 2018).

The quantitative component adopts a one-group pretest-posttest approach, in which participants are assessed before and after the six-month PBA program to describe changes in speaking performance. As this design does not include a control group and does not employ inferential statistical testing, the quantitative findings are interpreted descriptively, as an initial indication of change, rather than as statistical proof of a causal effect. The qualitative component consists of semi-structured interviews with selected students to explore their experiences and reflections, and is intended to help explain the patterns observed in the descriptive quantitative data. Students' perceptions are gathered using a closed-ended questionnaire, which is also analyzed descriptively. The qualitative data are intended to complement and contextualize the quantitative findings, providing a fuller understanding of how PBA was experienced in the compulsory extracurricular English Club setting at a boarding school.

Participants

The participants of this study are 40 students enrolled in a compulsory English Club program at a senior high school in a boarding school context in Indonesia. All 40 students participate in the quantitative phase, which includes the pretest, the six-month PBA treatment, the posttest, and the perception questionnaire. For the qualitative phase, 5 students were purposively selected after the posttest, based on their gain scores (the difference between pretest and posttest scores), to capture a range of experiences from students with the highest improvement to those with minimal or no improvement, while also considering representation across different proficiency levels (Creswell & Guetterman, 2018).

Research Instruments

Three instruments were used to collect data in this study, each corresponding to the research objectives. The primary instrument for measuring students' speaking ability was an analytic speaking rubric adapted from Brown (2004), as applied in Menggo and Gunas (2022). The rubric assesses student performance across five dimensions: comprehension, vocabulary, grammar, pronunciation, and fluency. Each dimension is scored on a four-point scale (1 = poor, 2 = fair, 3 = good, 4 = excellent), deliberately excluding a midpoint to encourage clear and decisive evaluative judgments. The analytic format was chosen because it allows each dimension to be evaluated independently, offering a detailed picture of student strengths and areas for development (Stanley, 2021b). The rubric was applied

THE IMPLEMENTATION OF PERFORMANCE-BASED ASSESSMENT
ON HIGH SCHOOL STUDENTS' SPEAKING ABILITY
IN AN EXTRACURRICULAR PROGRAM

consistently to both the pretest and posttest speaking interviews. Both tests were administered and scored by the same rater, the English teacher, to maintain consistency in scoring criteria across the two measurement points. This single-rater approach reduces variability in scoring standards but does not allow for inter-rater reliability estimation, which is acknowledged as a limitation of this study.

A closed-ended questionnaire was developed to measure students' perceptions of the PBA implementation in the English Club program. The questionnaire consisted of 8 items adapted from the perception instrument used in Menggo and Gunas (2022), which was originally grounded in the motivational framework of Schunk et al. (2008). The items covered key areas: confidence in speaking English, speaking anxiety, motivation to practice, perceived usefulness of tutor feedback, and overall perceived improvement in speaking ability. Each item was rated on a four-point Likert scale (1 = strongly disagree, 2 = disagree, 3 = agree, 4 = strongly agree), excluding a neutral midpoint to encourage decisive responses. The questionnaire was administered in Indonesian via Google Forms at the end of the program to ensure students could respond comfortably and honestly. Students were informed that their responses would not affect their scores and were anonymous.

A semi-structured interview guide was developed for selected students to gather their perspectives on the program. The interviews focused on students' experiences during the six-month PBA program, their perceptions of speaking improvement, challenges they faced, and their overall reflections on the PBA approach in the English Club context. The semi-structured format allowed for flexibility, enabling the researcher to probe deeper into emerging themes.

Research Instruments

The study was conducted over six months and followed three phases. Before program activities began, all 40 students completed an individual speaking pretest assessed using the analytic rubric. The pretest task required each student to speak on a given topic for a specified duration, providing a standardized baseline measure of speaking ability across the five rubric dimensions. Pretest scores were recorded and used to group students into beginner, intermediate, and advanced proficiency levels for the English Club program.

Following the pretest, students participated in the six-month English Club program, during which PBA was embedded throughout all speaking activities. Weekly meetings incorporated a range of communicative tasks aligned with CLT principles, including role plays, short presentations, and group discussions, as well as preparatory activities leading toward the end-of-term speaking performance (Richards, 2006). Throughout this phase, tutors applied the speaking rubric during performance activities and provided verbal feedback to students after each session. This ongoing assessment and feedback cycle reflects the process-oriented character of PBA, in which the value of assessment lies as much in supporting learning as in producing a final score (Stanley, 2021b). It is important to note that during regular classroom hours, students primarily focused on reading and writing exercises,

making the English Club the primary site for structured speaking practice and assessment.

At the end of the six-month period, all 40 students completed an individual speaking posttest, identical in format and assessment criteria to the pretest. The posttest was administered and scored by the same rater using the same analytic rubric to ensure consistency across measurement points (De Jong, 2023). Unlike the pretest, the posttest score was based solely on this final interview, not accumulated from daily scores, to maintain a direct and clean comparison between the two measurement points. Following the posttest, gain scores were calculated for all 40 students, and 5 students were selected for the interview phase based on their gain scores. Prior to the interviews, students were informed that participation was voluntary, that the interviews would be audio-recorded for transcription purposes, and that their identities would be kept confidential in all reporting. The student perception questionnaire was also distributed via Google Forms to all 40 participants at this stage. Students responded to 8 items based on their experience during the six-month program. The questionnaire was anonymous and took approximately 5–10 minutes to complete.

FINDING AND DISCUSSION

Finding

Speaking Score Results

Prior to the PBA program, a speaking pretest was administered to all 40 students using an analytic rubric adapted from Brown (2004), covering five dimensions: Comprehension, Vocabulary, Grammar, Pronunciation, and Fluency, each scored on a 1–4 scale.

Figure 1. Students' Speaking Scores Prior to PBA Implementation

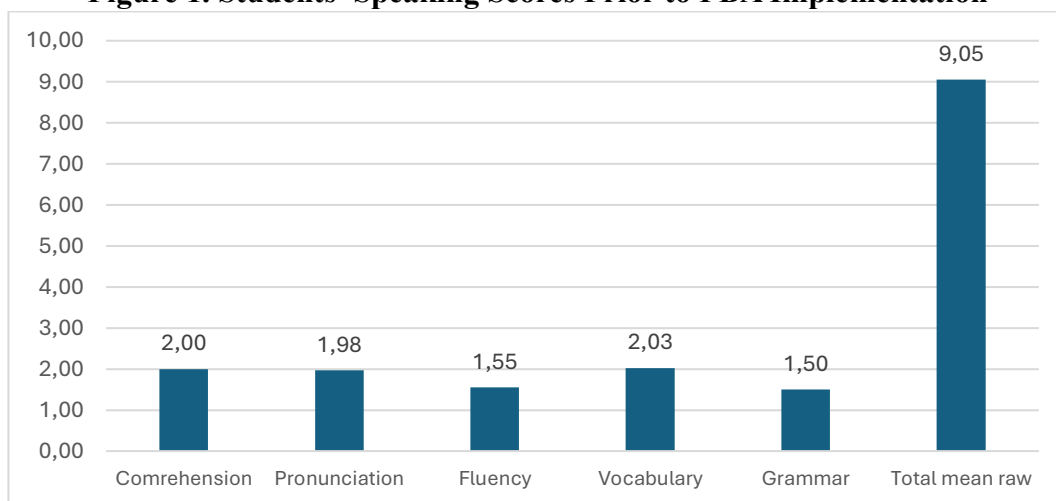


Figure 1 presents the mean pretest scores for each dimension. The total mean raw score was 9.05 out of a maximum of 20, placing students' initial speaking

ability in the "poor" to "fair" category. Among the five dimensions, Vocabulary obtained the highest mean (2.03), followed by Comprehension (2.00) and Pronunciation (1.98). Fluency (1.55) and Grammar (1.50) were the two weakest dimensions, indicating that students frequently experienced long pauses in spontaneous speech and made persistent grammatical errors in oral production. These results confirm that prior to the PBA intervention, students' speaking ability required substantial improvement across all dimensions, with fluency and grammar as the most critical areas.

After the six-month PBA program, a posttest was administered using the same rubric and procedures.

Figure 2. Students' Speaking Scores Following PBA Implementation

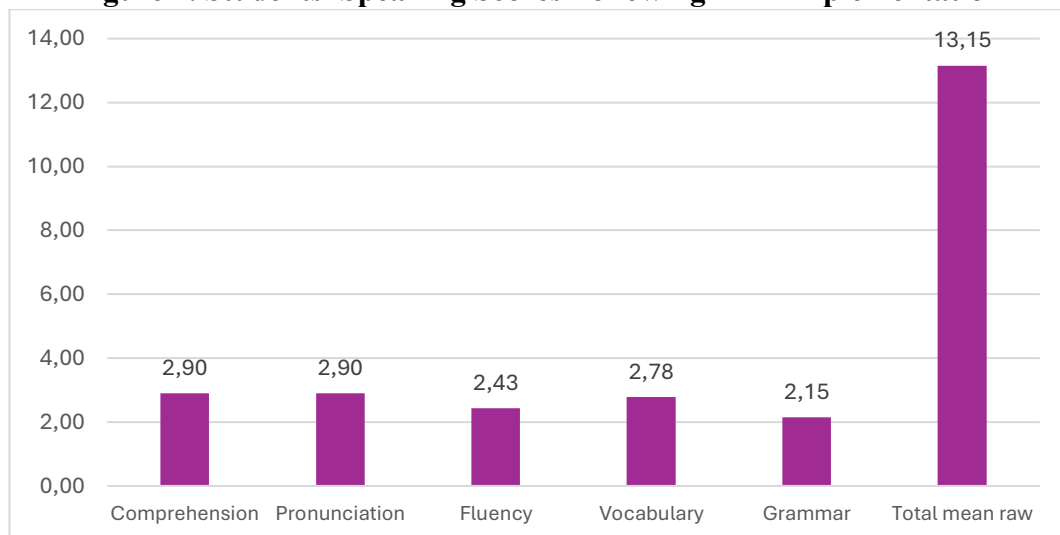
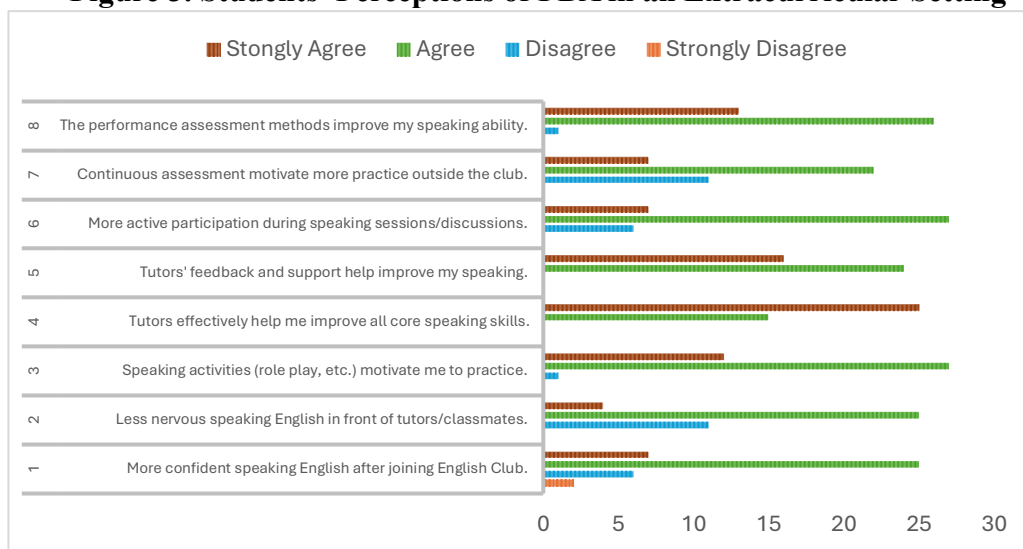


Figure 2 presents the posttest mean scores. The total mean raw score increased to 13.15. Comprehension and Pronunciation obtained the highest posttest means, both at 2.90, followed by Vocabulary (2.78), Fluency (2.43), and Grammar (2.15). The comparison of pretest and posttest scores showed a consistent descriptive increase across all five dimensions. Grammar, while showing the largest absolute gain from 1.50 to 2.15, remained the lowest-scoring dimension at posttest. As this study did not include a control group or inferential statistical testing, the observed score increase is reported descriptively and should not be interpreted as evidence of a causal effect of PBA on students' speaking ability.

Students' Perceptions of PBA

At the end of the program, a closed-ended questionnaire was administered to all 40 participants to capture their perceptions of the PBA approach.

Figure 3. Students' Perceptions of PBA in an Extracurricular Setting

The results indicate that students generally held positive perceptions of the program, with an overall mean score of 3.16, falling within the "Agree" category. The highest-scoring items were Item 4 (3.63), which measured students' perception that tutors effectively helped improve their core speaking skills, and Item 5 (3.40), which affirmed the value of tutor feedback and support. Item 8 (3.30) and Item 3 (3.28) also received strong agreement, reflecting positive views of the assessment method and speaking activities such as role plays and presentations. In contrast, the three lowest-scoring items concerned confidence in speaking (Item 1, 2.93), motivation to practice independently outside the club (Item 7, 2.90), and reduction of speaking anxiety (Item 2, 2.83). While all items fell within the "Agree" range, the comparatively weaker endorsement of these items suggests that confidence gains and anxiety reduction were present but unevenly distributed across participants.

Qualitative Findings

To gain a deeper understanding of students' experiences with the PBA program, semi-structured interviews were conducted with five purposively selected students based on their speaking gain scores, representing a range of improvement profiles. Table 1 presents the profiles of the five interviewees.

Table 1. Profiles of Interview Participants Based on Speaking Gain Scores

Student	Level	Pretest	Posttest	Gain	Selection Rationale
Student 1	Intermediate	55	95	+40	Highest gain
Student 2	Advanced	80	95	+15	Highest posttest score
Student 3	Intermediate	62	75	+13	Typical gain (close to mean)
Student 4	Beginner	50	60	+10	Low gain
Student 5	Beginner	45	45	0	No improvement

Theme 1: Initial Anxiety and Compulsory Participation

Nearly all participants reported considerable anxiety at the start of the program, particularly around the speaking pretest. Student 4 recalled: "I was really embarrassed. Afraid of being wrong. Afraid my answers were bad. Afraid my grammar was wrong. Afraid I wouldn't understand the questions." The intensity of this anxiety varied across participants, though it was widely shared. One student additionally linked initial reluctance to the compulsory nature of the program; this was raised by only one participant and is treated as an individual observation rather than a group-wide pattern.

Theme 2: Repeated Performance Increased Confidence

For students with higher speaking gains, initial anxiety gradually gave way to increased confidence through repeated exposure to performance tasks. Student 1 reflected: "Later on, especially once I had performed often, it became normal. I used to be scared to speak alone, especially in front of people, but now I'm quite brave, fluent." This pattern was consistently described by Students 1, 2, and 3. Students 4 and 5, who showed lower or no gains, did not report a comparable shift, a contrast examined further in Theme 5.

Theme 3: Constructive Tutor Feedback Guided Improvement

Positive perceptions of tutor feedback were expressed by all five students regardless of their speaking outcomes. Student 1 described the feedback as: "[The tutor] always gives comments... It's supportive, it motivates me... it's not criticism. It's like, 'you're already good, try fixing this part.'" Even Student 5, whose score did not improve, acknowledged that post-performance feedback was useful in identifying specific weaknesses. The universality of this response across all proficiency levels suggests that constructive feedback delivery was a consistent strength of the program.

Theme 4: Different Learning Responses Outside the Classroom

Independent practice outside the English Club varied considerably across participants. Student 1 described a self-directed routine involving self-recording, vocabulary review, and peer conversation. In contrast, Student 4 described practice as largely assignment-dependent: "Only if there's an assignment. I look up words on Google Translate... if there's no assignment, I usually don't practice." Student 5 reported minimal independent engagement with little retention of new language. This variation in self-regulated learning behavior broadly corresponded with the distribution of speaking gains observed across the five participants.

Theme 5: Challenges Limiting Speaking Improvement

For students with lower or no gains, the central challenge was not a failure to understand feedback but an inability to translate it into concrete practice. Student 5 explained: "I already know I should practice, but when I actually try to start, I

don't know where to begin. So in the end, I just don't." Student 4 similarly identified persistent vocabulary gaps as a recurring barrier during speaking tasks. These accounts suggest that for lower-proficiency students, the gap between understanding feedback and acting on it was a key limiting factor, pointing to the need for more structured guidance on how to practice between sessions.

Discussion

This study examined the implementation of Performance-Based Assessment in a compulsory extracurricular English Club at a boarding school. The discussion addresses each research question in turn, drawing on both quantitative and qualitative findings. The quantitative findings indicate a descriptive increase in students' mean speaking scores following the six-month PBA program, alongside broadly positive student perceptions and varied qualitative experiences across different proficiency levels.

Changes in Speaking Scores Following PBA Implementation

The total mean raw score rose from 9.05 at pretest to 13.15 at posttest out of a maximum of 20. This pattern is broadly consistent with previous studies. Menggo and Gunas (2022) documented a comparable increase from 60 to 76 following a three-month PBA intervention at the college level. Plengkham and Wasanasomsithi (2023) similarly found that posttest scores exceeded pretest scores across all communication modes in a PBA-integrated program for Thai undergraduates. The present findings align with this broader pattern, suggesting that structured speaking programs incorporating PBA principles, rubric-based feedback, and repeated communicative performance tasks may be associated with observable improvements in oral performance scores.

These findings must be interpreted cautiously. The study employs a one-group pretest-posttest design without a control group and does not include inferential statistical testing. Alternative explanations, including general language maturation over six months, the repeated testing effect, and students' independent learning activities outside the club, cannot be ruled out, as noted by Fadila and Trisno (2025). The findings therefore serve as an initial descriptive indication that PBA-embedded speaking programs may support score improvement, rather than as evidence of a definitive causal effect.

Among the five dimensions, Grammar remained the lowest-scoring at posttest despite showing the largest absolute gain from 1.50 to 2.15. This echoes Menggo and Gunas (2022), who also observed grammar as among the dimensions most resistant to short-term improvement in speaking contexts. This pattern suggests that while communicative exposure through PBA activities can support receptive comprehension and phonological production relatively quickly, the accurate application of grammatical structures in spontaneous oral production requires more sustained and targeted instructional attention, a consideration worth addressing in future implementations of PBA in similar extracurricular settings.

Students' Perceptions of PBA Implementation

The questionnaire results indicate that students generally perceived the PBA program positively, with an overall mean of 3.16 on a four-point scale. This is consistent with Menggo and Gunas (2022), who reported similarly positive questionnaire responses following a PBA speaking course, and with Plengkham and Wasanasomsithi (2023), who documented favorable student attitudes toward performance-based speaking tasks across multiple communication modes.

The highest endorsement was found for items related to tutor support and feedback, with Item 4 (3.63) and Item 5 (3.40) receiving the strongest agreement. When students can clearly identify what to improve and receive encouragement alongside correction, the assessment process becomes pedagogically meaningful rather than merely evaluative, a characteristic aligned with the process-oriented nature of PBA as described by Stanley (2021b), in which the value of assessment lies as much in the learning it generates as in the scores it produces.

The comparatively lower scores for confidence (Item 1, 2.93), motivation to practice independently (Item 7, 2.90), and reduction in speaking anxiety (Item 2, 2.83) suggest that PBA, as implemented in this program, was more effective in supporting structured in-program learning than in fostering autonomous practice or reliably reducing anxiety across all participants. This partially contrasts with Menggo and Gunas (2022), who reported stronger affective gains, including notable reductions in speaking anxiety. The difference may partly be explained by the compulsory nature of the present program and the wider range of student proficiency levels included.

Students' Experiences and Reflections on the PBA Program

The qualitative findings provide a richer account of how PBA was experienced across different proficiency and improvement profiles, and help explain several patterns observed in the quantitative data.

Regarding initial anxiety, nearly all participants reported considerable apprehension before and during the pretest, a high affective filter widely documented in speaking assessment contexts (Krashen, 1982). For students with higher gains (Students 1, 2, and 3), this anxiety diminished over time through repeated exposure to performance tasks, consistent with the principle that frequent low-stakes performance builds communicative automaticity over time (Richards, 2006). Students 4 and 5, however, reported that fear and hesitation persisted throughout the full six months of the program. This divergence suggests that PBA in a compulsory, mixed-proficiency setting may be highly effective for students who enter with moderate confidence but may require additional differentiated support for students whose anxiety is more deeply rooted.

Constructive tutor feedback was valued positively by all five students regardless of their speaking outcomes, reinforcing the questionnaire findings and confirming that the quality of formative feedback delivery was a consistent strength of the implementation. This aligns with Nadolski et al. (2021), who demonstrated that rubric-based formative feedback in speaking contexts is more effective than

global rating approaches in directing students toward specific areas for improvement. That even Student 5, who showed no score gain, acknowledged feedback as helpful points to the robustness of this element across proficiency levels.

A notable variation was observed in how students engaged with independent practice outside the club. Students with higher gains described varied and self-initiated routines, while students with lower gains reported practice that was minimal or contingent solely on assignment deadlines. This extends the findings of Menggo and Gunas (2022), who documented PBA's positive impact on study habits in a more structured college-level context, by suggesting that these motivational benefits may not transfer automatically to all learners in a compulsory secondary school extracurricular setting. For future implementations, incorporating brief guided sessions on self-directed practice strategies into the program itself may help students with lower motivation develop habits that higher-gaining students appeared to develop independently.

More critically, the accounts of Students 4 and 5 revealed that the central barrier to improvement was not a failure to understand feedback but an inability to translate it into concrete practice. Both students described knowing what needed to be improved while simultaneously not knowing how to begin. This gap between feedback comprehension and feedback application suggests that verbal feedback, while clearly delivered, may need to be accompanied by structured practice guidance, for instance, brief written practice plans specifying which dimension to focus on and suggesting concrete activities aligned with the student's proficiency level.

CONCLUSION AND SUGGESTION

This study examined the implementation of Performance-Based Assessment in a compulsory extracurricular English Club at a boarding school, combining quantitative score data with student perception and interview findings. The descriptive results indicate an increase in speaking scores across all five rubric dimensions following the six-month program, with the strongest gains observed in comprehension and pronunciation, while grammar remained the most persistent challenge. Students generally perceived the program positively, particularly in relation to tutor feedback and support, though gains in confidence, anxiety reduction, and independent practice outside the club were less consistently reported, especially among lower-proficiency participants. The qualitative findings further revealed that for some students, the central barrier to improvement was not a lack of feedback but an inability to translate that feedback into actionable practice, pointing to a gap between formative assessment and guided self-directed learning that future PBA implementations should address.

Several limitations of this study should be noted. The one-group pretest-posttest design without a control group means that the observed score increases cannot be attributed to PBA with confidence, as alternative explanations including

language maturation and the repeated testing effect cannot be ruled out. The use of a single rater and the absence of inferential statistical testing further limit the strength of the quantitative findings. The qualitative sample of five students, while purposively selected, is small and may not represent the full range of student experiences within the group. The compulsory boarding school context also limits the generalizability of the findings to other settings, particularly voluntary English Club programs in public schools. Future research should employ quasi-experimental designs with control groups, multiple raters, and larger qualitative samples to extend and strengthen these findings across more varied extracurricular contexts.

REFERENCES

- Afifah, E. S., & Widyantoro, A. (2025). HOTS Performance Based Assessment to improve Critical Thinking Skills in Speaking Class at ANNIDA Senior High School. *International Journal of Multicultural and Multireligious Understanding*, 12(10), 232. <https://doi.org/10.18415/ijmmu.v12i10.7103>
- Aprianto, D., & Sutarman, S. (2025). Impacts of Different Learning Intensities on Kinesthetic Efl Students' Speaking Performance. *Wiralodra English Journal*, 9(1), 77–90. <https://doi.org/10.31943/wej.v9i1.374>
- Bachman, L. F., & Palmer, a. S. A. S. (1996). Language Testing in Practice: Designing and Developing Useful Language Tests. In *Oxford applied linguistics*. (pp. vi, 377 p.).
- Bashori, M., van Hout, R., Strik, H., & Cucchiarini, C. (2024). I Can Speak: improving English pronunciation through automatic speech recognition-based language learning systems. *Innovation in Language Learning and Teaching*, 18(5), 443–461. <https://doi.org/10.1080/17501229.2024.2315101>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology Using thematic analysis in psychology. In *Psychiatric Quarterly* (Vol. 0887, Issue 1). <http://www.ncbi.nlm.nih.gov/pubmed/11752478>
- Brookhart, S. M. (2018). Appropriate Criteria: Key to Effective Rubrics. *Frontiers in Education*, 3(April). <https://doi.org/10.3389/educ.2018.00022>
- Brown, H. D. (2004). *Language Assessment: Principles and Classroom Practices*. Longman.
- Creswell, J. W., & Guetterman, T. C. (2018). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research* (6th Editio). Pearson.
- Creswell, J. W., & Plano Clark, V. L. (2017). Designing and Conducting Mixed Methods Research. In *Oxford Guides to Chaucer*. SAGE Publications, Inc. <https://doi.org/10.1093/oso/9780198821427.002.0002>
- De Jong, N. H. (2023). Assessing Second Language Speaking Proficiency. *Annual Review of Linguistics*, 9, 541–560. <https://doi.org/10.1146/annurev-linguistics-030521-052114>
- Dimova, S. (2022). Performance-based Speaking Tests: Possibilities in Local Language Testing. *Language Teaching Research Quarterly*, 29, 120–133.

- <https://doi.org/10.32038/ltrq.2022.29.08>
- Fadila, W., & Trisno, E. (2025). Enhancing Speaking Skills through Performance Assessment: A Study in Junior High School. *AL-ISHLAH: Jurnal Pendidikan*, 17(1), 1550–1562. <https://doi.org/10.35445/alishlah.v17i1.6255>
- Gisma, R. R., Nur, S., & Baa, S. (2025). Facilitating EFL Speaking Skills through an English Club. ... *English Language Teaching*, 6(4), 437–450. <https://www.foster.pbiiainpalopo.ac.id/index.php/contents/article/view/322%0Ahttps://www.foster.pbiiainpalopo.ac.id/index.php/contents/article/download/322/153>
- Krashen, S. D. (1982). *Principles and Practice in Second Language Acquisition* (1st ed.). Pergamon Press.
- Lestari, T., & Azizah, D. M. (2021). Implementing Performance-Based Formative Assessment in Grammar Class During Remote Teaching. *Wiralodra English Journal*, 5(1), 45–55. <http://wej.unwir.ac.id/index.php/wej/article/view/128%0Ahttp://wej.unwir.ac.id/index.php/wej/article/download/128/68>
- Maros, M. A., Ilma, R., Kurniasari, R., Program, E. S., Islam, U., Raden, N., & Palembang, F. (2023). *Speaking Activities Implemented In An English Club: Students' Voices*. 10(1), 81–92.
- Menggo, S., & Gunas, T. (2022). College Student's Perception of Performance-Based Assessment Use in Boosting Speaking Ability. *International Journal of Language Education*, 6(4), 423–436. <https://doi.org/10.26858/ijole.v6i4.22910>
- Nadolski, R. J., Hummel, H. G. K., Rusman, E., & Ackermans, K. (2021). Rubric formats for the formative assessment of oral presentation skills acquisition in secondary education. *Educational Technology Research and Development*, 69(5), 2663–2682. <https://doi.org/10.1007/s11423-021-10030-7>
- Nggawu, L. O., & Thao, N. T. P. (2023). The impact of Communicative Language Teaching (CLT) Approach on Students' Speaking Ability in a Public Indonesian University: Comparison between Introverts and Extrovert Groups. *International Journal of Language Education*, 7(3), 393–413. <https://doi.org/10.26858/ijole.v7i3.50617>
- Nunan, D. (1991). *Language teaching methodology: a textbook for teachers*.
- O Malley, J. M., & Valdez Pierce, L. (1996). *Authentic Assessment for English Language Learners: Practical Approaches for Teachers*. Addison-Wesley.
- Octaberlina, L. R., & Muslimin, A. I. (2022). Fostering Students' Speaking Ability through English Club Activities. *Arab World English Journal*, 13(3), 414–428. <https://doi.org/10.24093/awej/vol13no3.27>
- Plengkham, B., & Wasanasomsithi, P. (2023). Effects of Integrated Performance Assessment Modules on English Speaking Ability of Thai EFL Undergraduate Students. *LEARN Journal: Language Education and Acquisition Research Network*, 16(1), 448–472.
- Pratiwi, W. R. (2024). Evaluating the Impact of School-based English Clubs on English Language Acquisition and Learners' Self-confidence. *Journal of*

THE IMPLEMENTATION OF PERFORMANCE-BASED ASSESSMENT
ON HIGH SCHOOL STUDENTS' SPEAKING ABILITY
IN AN EXTRACURRICULAR PROGRAM

- Research on Language Education (JoRLE)*, 5(2), 1–9.
<https://ejurnal.teknokrat.ac.id/index.php/JoRLE/index>
- Putri, E., Ayu, D., Mujiyanto, J., & Wahyuni, S. (2024). *Implementation of Project-Based Assessment to Assess the Students' Speaking Skills in Vocational High School*. 14(1), 127–137.
- Richards, J. (2006). Communicative Language Teaching Today. In *Cambridge University Press* (Vol. 25, Issue 2).
- Salsabila, M., Damanik, I. J., & Girsang, S. E. E. (2025). The Influence of Joining English Club Toward Students Speaking Ability at 8TH Grade in UPTD SMP Negeri 6 Pematang Siantar. *Bilingual : Jurnal Pendidikan Bahasa Inggris*, 7(2), 176–183. <https://doi.org/10.36985/w8d0mp02>
- Schunk, D. H., Pintrich, P. R., & Meece, J. L. (2008). Motivation in Education: Theory, Research, and Applications. In *British Library Cataloguing-in-Publication Data* (Fourth). Pearson.
- Simbolon, H., Tampubolon, S., & Pasaribu, A. (2023). The Effect of Taking English Club Extracurricular Activity on Student's Writing and Speaking Ability. *IDEAS: Journal on English Language Teaching and Learning, Linguistics and Literature*, 11(1), 1–12.
<https://doi.org/10.24256/ideas.v11i1.3258>
- Stanley, T. (2021a). *Performance-Based Assessment for 21st-Century Skills*. Prufrock Press Inc.
- Stanley, T. (2021b). *Using Rubrics for Performance-Based Assessment*. Prufrock Press Inc.
- Zaim, M., Refnaldi, & 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.
<https://doi.org/10.29333/iji.2020.13340a>