



Self-Assessment in Islamic Education: Its Effectiveness on Speaking Performance at Madrasah Aliyah Al-Khoirot Malang

Mohammad Fiqri Fahriza Aswali^{1*}, Budik Kusworo²

^{1,2} Universitas Al-Qolam Malang, Indonesia

Mohammadfiqrifahriza21@alqolam.ac.id^{1*}, budikkusworo@alqolam.ac.id²

Korespondensi penulis: Mohammadfiqrifahriza21@alqolam.ac.id*

Abstract. This quasi-experimental study investigates the effectiveness of self-assessment in the context of Islamic education, specifically focusing on its impact on students' English speaking performance at Madrasah Aliyah Al-Khoirot Malang. The research adopts a pre-test and post-test control group design to measure students' development in speaking skills. A total of 40 secondary-level students participated, divided equally into experimental and control groups. The experimental group received instruction that incorporated self-assessment strategies, while the control group was taught using conventional methods without self-assessment integration. Data collection was conducted through structured speaking assessments administered before and after the intervention. The results demonstrated a notable improvement in the experimental group's speaking performance, with their mean score increasing from 57.65 (pre-test) to 63.90 (post-test). In contrast, the control group exhibited only a slight increase from 54.70 to 55.80. Statistical analysis using SPSS version 25 and a significance level of $\alpha = 0.05$ revealed a p -value of 0.039, which is less than 0.05, indicating a statistically significant difference between the two groups. These results confirm that self-assessment positively influences students' speaking proficiency. Beyond score improvements, students in the experimental group also showed increased self-awareness, better self-regulation, and deeper reflection on their speaking abilities, identifying their strengths and areas for improvement more effectively. This suggests that self-assessment fosters not only language performance but also metacognitive skills that are crucial for lifelong learning. This study contributes to the relatively limited body of research on self-assessment practices in Islamic education environments, particularly within madrasah contexts. It highlights the pedagogical value of self-assessment in enhancing speaking skills and suggests that incorporating self-reflective practices in language instruction can lead to more autonomous and engaged learners.

Keywords: Islamic Education, Language Learning, Learner Autonomy, Self-Assessment, Speaking Performance

1. INTRODUCTION

In the context of English language Teaching (ELT) in schools, the enhancement of spoken language proficiency is a significant challenge. It is common to observe a variation in students' speaking abilities, attributable to their differing levels of exposure and interaction with the second language, which exhibits variability due to sporadic contact. Several factors can influence pronunciation, including the influence of students' regional language or mother tongue backgrounds (Arjulayana & Martínez, 2022) shows that 40% of high school students' pronunciation errors stem from regional language interference. Furthermore, speaking skills are a publicly exhibited ability, is a prevalent issue, as evidenced by research in the domain of speaking skills (Aziz & Kashinathan, 2021). Consequently, it can be concluded that the development of speaking skills represents a significant challenge in an academic setting.

Given these challenges, empowering students through self-assessment techniques could play a pivotal role in improving spoken English proficiency. Research suggests that self-assessment fosters metacognitive awareness, allowing learners to identify their own

pronunciation errors, monitor progress, and reduce speaking anxiety (Ali Ahmed Qasem, 2020). For instance, structured self-assessment activities where students evaluate each other's speaking performances using guided rubrics can enhance self-awareness and collaborative learning (Kumar et al., 2023). By analysing peers' pronunciation, fluency, and body language, learners gain critical listening skills while reflecting on their own weaknesses (Abd Al Galil & Abd Al Galil, 2019). This method not only reduces reliance on teacher feedback but also builds a supportive classroom environment (Namaziandost et al., 2024). In contexts where regional language interference is prevalent such as 45FDZX in Indonesia guided self-assessment may help learners consciously recognize and correct L1-influenced errors (Alhammad, 2022). Further investigation is needed to explore how structured self-assessment protocols, tailored to diverse proficiency levels, can complement traditional teaching methods in mitigating the abovementioned challenges.

In the context of high school education, Self-Assessment is a method that remains underutilized in many schools. Huang (2016) state in his study that Self-Assessment has the potential to contribute to the enhancement of speaking skills by enabling students to identify and address their individual deficiencies. Furthermore, Self-Assessment is facilitate a more profound comprehension of the objectives and purposes of past examinations, thereby promoting self-regulated learning (Putri et al., 2023). Self-Assessment can also foster students' awareness of the quality of their work and the criteria applied to evaluate their work (Khabbzbashi, 2017).

Regarding the effectiveness of Self-Assessment method in English speaking performance, several studies have been conducted by Alghanmi (2023), in his research it is mentioned that Self-Assessment fosters a deeper level of reflection, where students analyze their own speaking performance, identify areas for improvement, and make informed decisions about how to improve their speaking skills. In addition, it makes them think about their weaknesses and strengths. Another study conducted by Quito Parra (2021), claims to have obtained positive results with students on their speaking skills after implementing the Self-Assessment technique. Another study by Borja Guzmán (2022), states that Self-Assessment helps students become autonomous and responsible for their tasks. Finally, Suquilanda Sigua (2022), mentions that the Self-Assessment technique helps students overcome difficulties in speaking skills and gives them confidence.

Numerous studies have found that incorporating Self-Assessment techniques enhances students' speaking abilities. Alfianti, (2022) research revealed that learners who engaged in Self-Assessment showed greater progress in oral proficiency than those evaluated through

traditional methods. Similarly, Andrade et al., (2021), a U.S.-based scholar, highlighted that formative Self-Assessment, particularly when students receive proper training, positively impacts both academic performance and self-regulated learning. Further supporting this, Vera Sánchez et al., (2023) conducted a study in Mexico, which found that Self-Assessment combined with self-regulation especially when facilitated via WhatsApp led to notable gains in fluency and better cognitive control over speech.

In writing this research, the author found a limited exploration of self-assessment practices within the context of madrasah level. Previous studies on self-assessment, particularly in the domain of speaking skills, have been conducted in Indonesia, with a specific focus on madrasah institutions. For instance, research by Purmanah et al., (2017) examined the application of self-assessment in the teaching of social studies (IPS) at MTs Sabilul Chalim. Similarly, studies at the Madrasah Aliyah level have also been conducted by Nu'man (2019), which highlighted the challenges associated with mathematics education, often perceived as difficult and intimidating. This study emphasized the importance of addressing psychological aspects, including self-awareness, to enhance learning outcomes. Furthermore, Arif (2021) explored the implementation of self-assessment in Islamic Religious Education (PAI) at Madrasah Aliyah Pondok Pesantren Daarun Nahdhah Thawalib Bangkinang. Collectively, these studies underscore the growing recognition of self-assessment as a valuable tool in madrasah education, yet they also reveal a need for further research to deepen understanding and expand its application across various subjects and educational levels particularly in the area of speaking skills in English.

From the various previous studies above, the originality of this study lies in the effectiveness of the self-assessment method on the speaking skills of grade 11 MIPA at MA Al-Khoirot Malang, which fills the gap in enriching literature of self-assessment in English learning at the Madrasah Aliyah level. Unlike previous research which focused more on subjects such as science (MIPA), Islamic education (PAI), and mathematics, this research pays special attention to English, especially in the aspect of students speaking performance, which is an aspect that often requires a comprehensive approach, where speaking skills are skills that have a higher level. With a comprehensive approach, the purpose of this study is for determine the effectiveness of the self-assessment method in English learning, especially in the aspect of students' speaking skills.

The purpose of this study is to analysis the speaking skills of Second grader senior high school from MA Al-Khoirot Putra by implementing the Self-Assessment technique. The researcher applied this method to the recruits' English material in class that use for

conversations. Furthermore, the researcher seeks the effectiveness of using the Self-Assessment technique can influence students' speaking performance. The first step of this study is pre-test before applying the treatment, determine the student's initial competence. what Self-Assessment is, how to use it with students, and its benefits and restrictions. It subsequently continues on to explain how the effectiveness happened in the classroom. It refers to how the students evaluated themselves and how the classes were organized. The researcher then continues on to explain the approach taken to conduct this study. This section describes the tools used to gather student data, the participants involved, the methods for analysing the data, and the ethical measures taken to protect the participants.

Next, the study presents its findings, including quantitative results and their interpretations. The researcher then discusses these outcomes, comparing and contrasting them with previous studies. Based on the analysis, final conclusions are drawn. The paper concludes by acknowledging the study's limitations and suggesting areas for future research. References and appendices are included at the end of the document.

2. METHOD

This study employs a quantitative methodology, utilizing a quasi-experimental research design to systematically evaluate the efficacy of Self-Assessment techniques in enhancing students' oral proficiency. The sample population comprises two intact classes of 11th-grade MIPA students (N=40), with Class A (n=20) designated as the experimental group exposed to the intervention, while Class B (n=20) serves as the control group under conventional assessment conditions. To triangulate the quantitative findings and measure learning gains, pre-test and post-test instruments were administered, generating comparative performance metrics to substantiate the research outcomes.

The target population of this study encompasses all 11th-grade students enrolled in the Natural Sciences (MIPA) program at Madrasah Aliyah Al-Khoirot Malang. A purposive sampling technique was applied, selecting two intact classes (11th MIPA A and 11th MIPA B) as the sample, yielding a total of 40 participants. These classes were strategically assigned into two distinct groups: Experimental group (11th MIPA A, n=20): Received the Self-Assessment intervention Control group (11th MIPA B, n=20): Followed traditional assessment methods This division ensures a comparative framework to evaluate the effectiveness of the implemented pedagogical approach.

The research instruments used in this study include a Survey this survey is designed to help researchers identify specific phonological interference patterns from Bahasa Indonesia

and regional languages that negatively impact Indonesian students' English speaking abilities. The data will reveal common pronunciation errors that hinder intelligibility and fluency, providing insights for targeted pedagogical interventions in ELT classrooms.

Speaking test to comprehensively assess the subjects' speaking development of the learning method. The speaking test was conducted individually in the form of a monologue, administered during both the pre-test and post-test phases to evaluate the significance of the subjects' progress in speaking skills. Questionnaire this questionnaire is designed to help researchers identify phonological interference patterns in Indonesian EFL learners that hinder English speaking proficiency, while also incorporating a self-assessment component to encourage learners to reflect on their pronunciation challenges. The instrument serves two key purposes: (1) diagnosing specific pronunciation errors stemming from L1 (Bahasa Indonesia or regional languages) interference, such as vowel/consonant substitutions, stress misplacement, and intonation issues; and (2) enabling learners to self-evaluate their speaking abilities, fostering metacognitive awareness of their strengths and weaknesses.

The questionnaire includes structured questions (Likert-scale self-ratings, multiple-choice error identification). By combining phonological error analysis with learner self-assessment, this tool not only helps researchers pinpoint common interference patterns but also empowers students to take an active role in improving their pronunciation. The results can inform targeted classroom interventions, such as minimal-pair drills or stress-timing exercises, while the self-assessment component promotes learner autonomy. This dual approach ensures a comprehensive understanding of phonological challenges while encouraging students to engage in reflective practice for long-term pronunciation improvement.

The quasi-experimental research design used is as follows:

O1 X O2
O3 X O4

Explanation:

O1: Pre-test of the experimental group O2: Post-test of the experimental group.

X: Application of the self-assessment method.

O3: Pre-test of the control group O4: Post-test of the control group.

The data collection techniques employed in this study include a speaking test designed to provide comprehensive data on students' speaking abilities and their interest in using the Self-Assessment method. The speaking test is conducted during both the pre-test and post-test phases, following identical procedures to ensure consistency, with the only difference being the timing of their application. The procedures for the speaking test are as follows: 1) the

researcher provides several topics including (*Talking about job, Daily activities, Describing people, Describing a process, Describing things*) for students to choose from randomly; 2) students are then given 3 minutes to speak on the selected topic; 3) the assessment is carried out simultaneously by the researcher and members of the research team to ensure reliability; and 4) the entire process is recorded to allow for subsequent review and verification of the data.

Data Analysis Technique

The data analysis technique in this study is adapted from (Heaton, 1991) analysis technique, which measures Pronunciation, Fluency, and Comprehensibility.

Table 1. Classification point from Heaton

Aspect	Poin	Classification	Criteria
Pronunciation	50	Very Good	Pronunciation is very clear and close to that of a native speaker.
	40	Good	Pronunciation is clear with few errors that do not interfere with meaning.
	30	Fair	Pronunciation is clear enough, but there are some errors that interfere with understanding.
	20	Poor	Pronunciation is often unclear and interferes with understanding.
	10	Very Poor	Pronunciation is very unclear and difficult to understand.
Fluency	50	Very Good	Fluency of speech is very good, almost like a native speaker.
	40	Good	Fluency is good with few pauses or repetitions.
	30	Fair	Fluency is sufficient, but there are frequent pauses or repetitions.
	20	Poor	Fluency is low with many pauses or repetitions.
	10	Very Poor	Fluency is very low and hinders communication.
Comprehensibility	50	Very Good	Very easy to understand, ideas are presented clearly and logically.

	40	Good	Easy to understand with few unobtrusive errors.
	30	Fair	Quite understandable, but there are some distracting errors.
	20	Poor	Difficult to understand because of many mistakes.
	10	Very Poor	Very difficult to understand because of severe mistakes.

The following formula is used to calculate students' test poin:

Conversion and Classification of Student poin

Speaking scores are converted using the following formula by (Sudijono, 2013):

$$\text{StudentPoin} = \frac{\text{the GainPoin}}{\text{the MaximumPoin}} \times 100$$

To systematically categorize the assessment outcomes, this study employs Heaton's (1990) classification framework, which provides standardized criteria for interpreting student performance levels. The classification scheme is structured as follows:

Table 2. Classification of student' Poin

No.	Poin	Classification	Criteria
1.	81-100	A	Very Good
2.	66-80	B	Good
3.	56-65	C	Fair
4.	41-55	D	Poor
5.	<-40	E	Very Poor

This study utilizes Heaton's established classification framework to ensure standardized evaluation of learner competencies. To determine intervention effectiveness, a paired samples t-test was conducted comparing pre-test and post-test results. Statistical analysis was performed using SPSS version 25 (IBM Corp), with an alpha level of 0.05 set as the threshold for statistical significance. This analytical approach enables rigorous comparison between experimental and control groups while maintaining measurement reliability.

3. FINDING AND DISCUSSION

Finding

This study presents a comprehensive analysis of pretest and post-test data to evaluate the development of students' speaking skills following the instructional intervention. The examination incorporates descriptive statistics including mean scores and standard deviations from both assessment phases, along with supplementary questionnaire responses, to provide a multidimensional evaluation of learning outcomes. Through systematic statistical analysis featuring frequency distributions, measures of central tendency (mean, median), and variability indices, the data reveals significant patterns in oral proficiency progression. These quantitative findings are further enriched by interpretive analysis that contextualizes the results within the framework of pedagogical effectiveness, offering substantive evidence of the intervention's impact while maintaining rigorous methodological standards through comprehensive data triangulation.

This section analyzes the initial speaking competence distribution across both experimental and control groups through frequency and percentage calculations. As Table 3 illustrates, students' pre-test scores were categorized into proficiency tiers based on established assessment criteria.

Pre-test speaking performance was classified to compare starting proficiency between experimental and control cohorts.

Table 3 presents the frequency distribution and percentage breakdown of students' pre-test speaking proficiency scores across experimental and control groups.

Criteria	Poin	Control Group		Experiment Group	
		N	Presentage	N	Presentage
Very Good	81-100	0	0	0	0
Good	66-80	3	15%	4	20%
Fair	56-65	5	25%	7	35%
Poor	41-55	8	40%	6	30%
Very Poor	<-40	4	20%	3	15%
Result		20	100%	20	100%

The baseline assessment revealed comparable speaking proficiency distributions between experimental and control groups. In the experimental cohort, 20% (n=4) of students demonstrated 'Good' proficiency, while 35% (n=7) achieved 'Fair' scores. The majority (30%,

n=6) fell into the 'Poor' category, with the remaining 15% (n=3) scoring 'Very Poor'. Parallel results emerged in the control group, though with slightly weaker performance: 15% (n=3) attained 'Good' scores, 25% (n=5) were rated 'Fair', while 40% (n=8) and 20% (n=4) placed in the 'Poor' and 'Very Poor' categories respectively.

The result found that the speaking proficiency levels of both groups were comparable, with the majority of students in each group being categorized as poor performers. This similarity in distribution implies that the initial speaking abilities of the two groups were statistically equivalent before the intervention.

Classification of students' post-test scores for experiment and control groups

Following pre-test administration, differentiated instructional approaches were implemented across the eight-week intervention period. The experimental group (n=20) received instruction incorporating structured self-feedback mechanisms, while the control group (n=20) continued with traditional teacher-led pedagogy. Both groups maintained identical contact hours across four instructional sessions, ensuring comparable time-on-task measurements.

Criteria	Poin	Control Group		Experiment Group	
		N	Presentage	N	Presentage
Very Good	81-100	0	0	0	0
Good	66-80	4	20%	9	45%
Fair	56-65	6	30%	8	40%
Poor	41-55	6	30%	2	10%
Very Poor	<-40	4	20%	1	5%
Result		20	100%	20	100%

The results presented in Table 4 reveal a significant disparity in post-test score distributions between the experimental and control groups. The experimental group demonstrated consistent competency improvement, with 9 participants (45%) classified as 'good,' an equal proportion 8 participants (40%) attaining 'fair,' 4 individuals (10%) scoring 'poor,' and the remaining 1 (5%) categorized as 'very poor.' This distribution suggests a relatively balanced progression in performance levels among the intervention group.

On the other side the control group exhibited a markedly different performance pattern. Only 4 participants (20%) met the 'good' poin, while 6 individuals (30%) achieved a 'fair' poin. The majority of the control cohort 6 participants (30%) remained at the 'poor' level,

supplemented by 4 subjects (20%) performing at a 'very poor' standard. This divergence in outcomes underscores the potential efficacy of the experimental intervention, as the control group's distribution skewed toward lower proficiency tiers, with fewer high achievers and a greater concentration of participants in the lower performance categories.

Mean value and standard deviation of students' pretest for the experiment group and control group.

After calculating the students' pre-test results, it shows that the control group is higher than the experimental group. This is shown in the table below:

Table 5 Mean Value and Standard Deviation of Student Pre-Test

Group	Average poin	Std. Deviation
Experemintal	57.65	14.40495
Control	54.70	13.56505

Table 5 showed that the mean pre-test score of the experimental group was 57.65, classified as low, while the control group obtained a mean score of 54.70, also falling into the low proficiency category. These results indicate that the initial academic performance of both groups was statistically comparable, with no significant difference in their baseline competencies. This parity suggests that the two groups exhibited similar learning productivity prior to the administration of the experimental treatment.

To evaluate whether the observed differences in post-test performance were statistically significant, an independent samples t-test was conducted using SPSS 25.

This parametric test was selected to compare the mean post-test scores between the experimental and control groups, assuming normal distribution and homogeneity of variance. The analysis aimed to determine whether the intervention yielded a measurable impact on student achievement beyond random variation.

Table 6 Independent Samples t-test for Student Pre-Test

Variabel	Probability Poin	Significance asymptotic
Pre-tes	0.05	.509

The table above shows that the statistical hypothesis is supported by an asymptotic statistical test. With a significance value (2-tailed) of less than .05 ($p = .000 < .05$), it can be concluded that H1 is accepted, while H0 is rejected. This indicates that student self-feedback

contributed significantly more to learning outcomes than the conventional method. Thus, it can be argued that self-feedback enhances students' speaking ability more effectively. These findings suggest that self-feedback should be considered a viable technique for improving English speaking skills among 11th-grade MIPA students at Al-Khoirot Islamic Senior High School (Madrasah Aliyah) in Malang, rather than being overlooked.

The Mean Score and Standard Deviation of Students' Post-Test

This section examines the differences in students' speaking ability following the post-test treatment. Initially, both the experimental and control groups had comparable mean scores, confirming their suitability for the study. The experimental group received instruction using student self-feedback, while the control group was taught through a conventional method focused on speaking skills. After the treatment, the students in both groups were given post-test to find out their speaking ability at the same level or not by using independent sample t-test analysed with SPSS 25. The findings of post-test are presented in table.

Table 7 Mean Value and Standard Deviation of Students' Post-Test

Group	Average poin	Std. Deviation
Experemintal	63.90	10.49260
Control	55.80	13.28077

The table above reveals a noticeable difference in the post-test mean scores between the two groups following the treatment. The experimental group achieved a mean score of 63.90, classified as fair, while the control group obtained a mean of 55.80, falling into the poor category ($63.90 > 55.80$). Additionally, the standard deviations were 10.49 for the experimental group and 13.28 for the control group. To determine whether this difference in post-test scores was statistically significant, an independent samples t-test was conducted using SPSS version 25.

Table 8 The Independent Sample T-Test of the Students' Post-Test

Group	Probability Poin	Significance asymptotic
Post-test	0.05	0.039

The statistical analysis presented in the table demonstrates that the hypothesis was tested using asymptotic statistical methods. The two-tailed significance value of .039 ($p < .05$) led to the rejection of the null hypothesis (H_0) and acceptance of the alternative hypothesis H_1 . This statistically significant result indicates that student self-feedback yielded substantially

better outcomes compared to conventional teaching methods. These findings clearly demonstrate that implementing student self-feedback effectively enhances students' speaking abilities.

Discussion

This discussion section interprets the pretest and post-test results, analyzing the collected data to compare speaking skill development between the experimental and control groups. The analysis reveals significant findings: The experimental class, which received the student self-assessment intervention, demonstrated notable improvement from a pretest average of 57.65 to a post-test average of 63.90. In contrast, the control group showed minimal progress, moving from 54.70 in the pretest to only 55.80 in the post-test.

This aligns with prior research Alghanmi, (2023); Andrade et al., (2021), which found that self-assessment fosters metacognitive awareness, reduces speaking anxiety, and promotes learner autonomy. The success of self-assessment can be attributed to its structured feedback mechanism. Students were trained to evaluate their own and peers' speaking performances, leading to heightened self-awareness and targeted improvements in pronunciation, fluency, and comprehensibility. This method not only encouraged reflective learning but also minimized reliance on teacher feedback, fostering a more autonomous learning environment (Kumar et al., 2023).

The post-test results demonstrate that implementing self-assessment methodology significantly enhances student achievement in speaking skills. This improvement is substantiated by the measurable increase in average scores between pre- and post-treatment assessments, as previously detailed. Comparative analysis reveals superior speaking skill development in the experimental class relative to the control group. Statistical validation through an independent samples t-test (SPSS Version 25) yielded a probability value ($p = 0.000$) significantly below the established alpha threshold ($\alpha = 0.05$). This statistically significant result confirms acceptance of the alternative hypothesis (H_1), supporting the effectiveness of the self-assessment intervention.

Self-assessment is an effective way to enhance students' speaking skills. They learn to evaluate their own performance, with teachers guiding them by having them observe and assess their peers before offering constructive feedback. This process of observing and critiquing others eventually helps students reflect on their own speaking abilities and identify areas for improvement. Through self-assessment, students can gauge how well they have grasped the

taught material. Additionally, this method allows them to refine their skills independently, serving as a form of self-correction and progress tracking in their learning journey.

4. CONCLUSION

This study provides empirical evidence that self-assessment significantly enhances English speaking performance among students at Madrasah Aliyah Al-Khoirot Malang. The experimental group, which underwent self-assessment-based instruction, demonstrated a marked improvement in speaking proficiency (mean post-test score: 63.90) compared to the control group (mean post-test score: 55.80). Statistical analysis ($p = 0.039 < 0.05$) confirmed the intervention's effectiveness, supporting the hypothesis that structured self-assessment fosters metacognitive awareness, self-regulation, and targeted skill improvement. These findings align a critical gap enriching the theory of self-assessment in EFL pedagogy within Islamic education by validating self-assessment as a viable strategy for improving oral communication skills in madrasah settings.

SUGGESTION

For Educators:

Integrate self-assessment training by providing explicit instruction on self-evaluation techniques, including rubrics for pronunciation, fluency, and comprehensibility. Encourage peer feedback by combining self-assessment with peer review to promote reflective learning and collaborative improvement. Reduce teacher dependence by gradually shifting from teacher-centered feedback to student-led evaluation, fostering learner autonomy.

For Curriculum Developers:

Design self-assessment modules that incorporate structured self-assessment frameworks into EFL curriculum, particularly for speaking skills. Adapt to local contexts by addressing regional language interference through targeted exercises.

For Future Research:

Conduct longitudinal studies to investigate the sustained impact of self-assessment on speaking retention and motivation over time. Explore technology integration by utilizing digital tools to facilitate self-assessment in hybrid learning environments. Broaden sampling to include diverse madrasah populations to enhance the generalizability of findings across Islamic education systems. By adopting these strategies, educators can empower students to take greater ownership of their learning, effectively bridging the gap between traditional instruction and autonomous skill development in EFL madrasah settings.

REFERENCE

- Abd Al Galil, H., & Abd Al Galil, M. (2019). *The effect of using reflective listening on developing EFL adults' oral fluency*. Online Submission.
- Alfianti, M. (2022). The implementation of self-assessment in EFL student's speaking performance. *Journal of English Language Teaching and Literature (JELTL)*, 5(2), 103–116. <https://doi.org/10.47080/jeltl.v5i2.2043>
- Alghanmi, B. F. (2023). *Self-assessment in EFL speaking classroom and its effect on achievement, self-regulated learning, and critical thinking: Students' voices from Saudi Arabia* [Doctoral dissertation, University of Glasgow].
- Alhammad, A. (2022). *The effect of explicit instruction on developing appropriate spoken pragmatic language choices in female Saudi Arabian English as a foreign language students* [Doctoral dissertation, The University of Liverpool (United Kingdom)].
- Ali Ahmed Qasem, F. (2020). The effective role of learners' self-assessment tasks in enhancing learning English as a second language. *Arab World English Journal*, 11(3), 502–514. <https://doi.org/10.24093/awej/vol11no3.33>
- Andrade, H. L., Brookhart, S. M., & Yu, E. C. (2021). Classroom assessment as co-regulated learning: A systematic review. *Frontiers in Education*, 6, 751168.
- Arif, M. (2021). *Implementasi self assessment dalam pembelajaran pendidikan agama Islam di Madrasah Aliyah Pondok Pesantren Daarun Nahdhah Thawalib Bangkinang*, 6.
- Arjulayana, A., & Martínez, M. M. (2022). Student's pronunciation error in speaking performance. *Globish: An English-Indonesian Journal for English, Education, and Culture*, 11(2), 101–108.
- Aziz, A. A., & Kashinathan, S. (2021). ESL learners' challenges in speaking English in Malaysian classroom. *Development*, 10(2), 983–991.
- Borja Guzmán, R. E. (2022). *Applying self-assessment to improve speaking skills in a secondary school* [Master's thesis, Universidad Casa Grande. Departamento de Posgrado].
- Heaton, J. B. (1991). *Speaking English language testing*. London: Longman.
- Huang, S.-C. (2016). Understanding learners' self-assessment and self-feedback on their foreign language speaking performance. *Assessment & Evaluation in Higher Education*, 41(6), 803–820.
- Ika Purmanah, N., Puspitasari, E., & Syekh Nurjati Cirebon, I. (2017). Penerapan self-assessment untuk menumbuhkan kesadaran siswa tentang makna belajar pada mata pelajaran IPS di MTs Sabilul Chalim Kecamatan Leuwimunding Kabupaten Majalengka. *Jurnal Edueksos*, 6(1), 65–80.
- Khabbazbashi, N. (2017). Topic and background knowledge effects on performance in speaking assessment. *Language Testing*, 34(1), 23–48.

- Kumar, T., Soozandehfar, S. M. A., Hashemifardnia, A., & Mombeini, R. (2023). Self vs. peer assessment activities in EFL-speaking classes: Impacts on students' self-regulated learning, critical thinking, and problem-solving skills. *Language Testing in Asia*, 13(1). <https://doi.org/10.1186/s40468-023-00251-3>
- Namaziandost, E., Kargar Behbahani, H., & Heydarnejad, T. (2024). Tapping the alphabets of learning-oriented assessment: Self-assessment, classroom climate, mindsets, trait emotional intelligence, and academic engagement are in focus. *Language Testing in Asia*, 14(1), 21.
- Nu'man, M. (2019). Self awareness siswa madrasah aliyah dalam pembelajaran matematika. *Jurnal Pengembangan Pembelajaran Matematika*, 1(1), 51–58. <https://doi.org/10.14421/jppm.2019.11.51-58>
- Putri, T. E., Parisyi, A., Salfen, H., & Sohiron, S. (2023). Peningkatan kinerja guru melalui implementasi self-assessment: Sebuah analisis terhadap dampaknya pada mutu pendidikan. *Didaktika: Jurnal Kependidikan*, 12(4), 911–920.
- Quito Parra, L. L. (2021). *Self-assessment of roleplays to improve speaking* [Master's thesis, Universidad Casa Grande. Departamento de Posgrado].
- Sudijono, A. (2013). *Pengantar evaluasi pendidikan*. Jakarta: RajaGrafindo Persada.
- Suquilanda Sigua, J. Y. (2022). *Self-assessment influence on improving speaking skills* [Master's thesis, Universidad Casa Grande. Departamento de Posgrado].
- Vera Sánchez, D. M., Pinargote Sánchez, M. J., & Lara Alcívar, D. K. (2023). Improving speaking skills by implementing self-assessment and self-regulation facilitated by WhatsApp in 8th year students. *Ciencia Latina Revista Científica Multidisciplinar*, 7(1), 2601–2619. https://doi.org/10.37811/cl_rcm.v7i1.4611