

ISSN 2503 - 3492 (Online)



JEES

(Journal of English Educators Society)

Volume 10 | No. 2 | Oktober 2025 | Sidoarjo

TABLE OF CONTENT

Cover

Table of Content

A Meta-analysis of the effects of differentiated instruction on English language proficiency	85
<i>Riza Weganofa, Mujiono</i>	
From classroom to screen: Vocational English students' copywriting practices in creating tourism promotion videos	95
<i>Dhini Aulia, Ahmad Bukhori Muslim, Budi Hermawan</i>	
Involving students in a collaborative and reflective reading material development: A participatory action research study	106
<i>Elok Putri Nimasari, Bambang Harmanto, Rinjani Bonavidi, Adi Fajaryanto Cobantoro, Elvina Dwi Cahyani, Velrinda Fahza'a Risty</i>	
Enhancing vocabulary acquisition through word association recitation among EFL students	118
<i>Andika Rikarno Putra, Dodi Mulyadi, Siti Aimah, Pirman Ginting, Charanjit Kaur Swaran Singh</i>	
From struggles to strategies: Indonesian students' use of e-portfolios in public speaking courses through the lens of self-regulated learning and sociocultural theory	135
<i>Raden Muhammad Ali, Pamastu Narpaduita, Muhammad Iqwan Sanjani, Dwi Hastuti, Ummi Khaerati Syam, Muhammad Kunta Biddinika</i>	
The influence of AI-Generated feedback on university students' debate skills	145
<i>Landry Dwiyoga Daniswara</i>	
Implementing and evaluating OpenAI whisper for accurate speaking assessment and skill development in Indonesian EFL classroom	156
<i>Salsabila Latifa, Nirwanto Maruf</i>	
Revealing learner beliefs in EFL learning across cultures and technologies	170
<i>Vega Hesmatantya</i>	
Enhancing EFL students' reading comprehension through Instagram: A quasi-experimental study	177
<i>Sudiran, Wahyu Taufiq, A.B. Prabowo Kusumo Adi</i>	

Enhancing students' speaking skills using ChatGPT application in ESP maritime English learning	186
<i>Dyah Ratnaningsih</i>	
Exploring peer review as a pedagogical tool to enhance paragraph-writing skills in university students	198
<i>Aidil Syah Putra, Abdul Rohim, Noor Azlan Ahmad Zanzali</i>	
Creative teaching in English language classrooms: A reflective and contextual interpretation	206
<i>Muhamad Ahsanu, Erna Wardani, Ashari Ashari, Nisa Roiyasa</i>	
Exploring the role of language policy in facilitating multilingual communication among pesantren students	219
<i>Alfiyatus Sa'adah, Zuliati Rohmah, Esti Junining, Hazlina Abdullah</i>	
Inclusive education: English teachers' challenges and strategies for hearing-impaired students in Indonesia	225
<i>Jati Suryanto, Berliana Farras Rachmawati, Saefurrohman, Maryam Sorohiti</i>	



A Meta-analysis of the effects of differentiated instruction on English language proficiency

Riza Weganofa,* Mujiono

Universitas PGRI Kanjuruhan Malang, Indonesia

This meta-analysis systematically examined the effects of differentiated instruction (DI) on English language proficiency, addressing the growing need for instructional approaches that address diverse student profiles. Seven empirical studies published between 2017 and 2024 were included. The overall pooled mean difference was 2.92 with a 95% confidence interval [1.64–4.19], indicating significant improvements in learners' grammar, reading comprehension, and fluency. A high heterogeneity level ($I^2 = 87\%$) was observed. The findings support DI as an effective pedagogical approach in multilingual English classrooms, consistently demonstrating its positive impact on grammar, reading comprehension, fluency, and overall language achievement compared to traditional approaches. However, substantial variability in effect sizes highlighted the influence of educational contexts, student characteristics, and DI implementation strategies. Methodological concerns, particularly regarding randomization and allocation concealment, also limited internal validity and generalizability. Despite these limitations, the findings reinforce DI's potential to promote equitable learning and enhance student engagement, advocating for its broader application in multilingual settings. Future research should adopt more rigorous experimental designs, longitudinal approaches, and comprehensive evaluations of implementation fidelity, while exploring contextual factors that shape DI effectiveness, to refine both theoretical frameworks and practical guidelines.

Keywords: differentiated instruction, effectiveness, language proficiency, meta-analysis

OPEN ACCESS

ISSN 2503 3492 (online)

*Correspondence:

Riza Weganofa

rizaweganofa@unikama.ac.id

Received: 12th May 2025

Accepted: 14th August 2025

Published: 2nd October 2025

Citation:

Weganofa, R., Mujiono, M. (2025). A Meta-analysis of the effects of differentiated instruction on English language proficiency. *JEES (Journal of English Educators Society)*, 10(2). <https://doi.org/10.21070/jees.v10i2.1929>

INTRODUCTION

In the field of language education, the growing diversity of student needs has intensified attention on instructional methodologies capable of addressing varied proficiencies, learning styles, and personal interests. Differentiated instruction (DI) has emerged as an innovative pedagogical framework that provides customized educational experiences, effectively responding to the specific needs of students in English language classrooms. Its efficacy is well supported by numerous studies highlighting substantial positive impacts on learners' academic outcomes and engagement across diverse contexts ([Hatmanto & Rahmawati, 2023](#); [Melka & Jatta, 2022a](#); [Ojong, 2023](#); [Zólyomi, 2022](#)). Given students' varied linguistic backgrounds and capabilities, integrating DI into educational practices creates a supportive environment crucial for enhancing English language proficiency, an indispensable skill in today's multilingual global landscape ([Ojong, 2023](#); [Sahril et al., 2021](#)).

The necessity of instructional differentiation becomes particularly critical in addressing the unique challenges encountered by language learners, such as inadequate language skills, varying motivational levels, and diverse cognitive capacities ([Ahmed, 2021](#); [Sari, 2018](#)). EFL classrooms are increasingly heterogenous comprising student who differ in prior knowledge, cultural backgrounds, and learning habits ([Mahali & Sevigny, 2022](#)).

Recognizing these distinct characteristics allows educators to employ strategies that not only support language acquisition but also foster inclusive and equitable educational opportunities (Ismail & Al Allaq, 2019; Zólyomi, 2022). Addressing these concerns is particularly urgent given the limitations of traditional, uniform instructional methods, which risk marginalizing students who struggle in standardized learning environments (Lawrence-Brown, 2020). Consequently, a thorough understanding of both the theoretical foundations and practical implications of DI is imperative for educators seeking meaningful improvements in English language teaching and learning outcomes (Hatmanto & Rahmawati, 2023; Ojong, 2023).

DI stems from constructivist learning theory, emphasizing learner-centered strategies that respond to student diversity (Ortega et al., 2018). This approach incorporates key constructs such as student variability, scaffolding, and formative assessment, which are essential differentiating content, processes, and products according to diverse student needs (Nassaji, 2017; Watts-Taffe et al., 2012). Empirical evidence further indicates that implementing DI frameworks encourages teachers to move beyond the traditional pedagogy, prompting learners to engage more deeply and fostering essential skills for lifelong learning (Sahril et al., 2021). Nonetheless, initial reluctance towards DI, often linked to misconceptions and insufficient professional training, remains a notable challenge to its broader adoption (Mavidou & Kakana, 2019).

Failure to adequately support language learners in overcoming disparities may have serious consequences. Students may find that developing core language skills is not merely challenging but essential for academic success, social interaction, and career advancement (Halil et al., 2024; Ojong, 2023). Insufficient differentiation can also lead students to experience disengagement and feelings of inadequacy, negatively affecting their educational experience and psychological well-being (Sahril et al., 2021). From a pragmatic standpoint, insufficient attention to student diversity creates substantial challenges for educators and policymakers striving to establish inclusive and effective learning environments (Hatmanto & Rahmawati, 2023; Ismail & Al Allaq, 2019). A lack of inclusivity in the classroom may leave some students vulnerable to underachievement, reduced motivation, disengagement, learning setbacks, and an inability to realize their full potential (Onyishi & Sefotho, 2020). Consequently, rigorous examination and refinement of DI methodologies are essential to ensure that all students benefit from high-quality education and can achieve their potential (Hatmanto & Rahmawati, 2023; Spencer-Waterman, 2014). Given these contextual considerations, this meta-analysis systematically evaluates existing research on the impact of DI on English language proficiency. Core methodological procedures include comprehensive literature reviews, careful analysis of research designs, and critical assessments of pedagogical effectiveness across diverse educational contexts (Melka & Jatta, 2022b; Sahril et al., 2021). This study principally aims to identify empirically supported DI strategies that enhance the effectiveness of English language teaching, while also highlighting gaps in the current research to provide a

conceptual foundation for future investigations and pedagogical advancements in this domain.

Despite extensive evidence demonstrating the beneficial effects of DI on language learning, notable gaps remain concerning its practical application across varied educational environments and diverse student populations (Sari, 2018; Zólyomi, 2022). Previous research often overlooks the complexities associated with teacher beliefs, institutional constraints, and the consistent implementation of DI strategies in real-world classrooms (Chien, 2015; Hatmanto & Rahmawati, 2023; Sari, 2018). Although DI has been increasingly adopted in language education, empirical findings remain fragmented and inconclusive, particularly within multilingual EFL contexts. Existing studies exhibit methodological inconsistencies and rarely translate their outcomes into actionable pedagogical implications. Moreover, a comprehensive meta-analytic synthesis focusing on DI's effectiveness in enhancing English proficiency has not yet been conducted. This study addresses these gaps by offering an evidence-based evaluation of DI's impact, methodological rigor, and implications for inclusive language instruction. By systematically synthesizing the existing literature and emphasizing practical applications, it seeks to advance an integrated comprehension of DI's role in promoting English language proficiency (Melka & Jatta, 2022a; Zólyomi, 2022). Ultimately, the purpose of this meta-analysis is to identify effective DI practices, offer theoretical contributions, and generate actionable recommendations for educators and stakeholders engaged in language education.

Through meticulous literature analysis and synthesis, this study aims to comprehensively investigate the effects of DI on English language proficiency and provide insights essential for improving instructional effectiveness. As a pedagogical framework that directly addresses learner diversity, DI holds strategic relevance for curriculum design and inclusive teaching practices. This rationale is established in the abstract and introduction, while the conclusion highlights broader implications and future directions for instructional innovation, avoiding unnecessary repetition of earlier justifications.

METHODS

Search strategy

A systematic literature search was conducted in alignment with PRISMA guidelines, focusing on peer-reviewed journal publications from 2017 to 2024. Electronic databases including Scopus, Crossref, and Google Scholar were searched using a combination of relevant keywords and Boolean operators such as differentiated instruction, English language proficiency, EFL, achievement, and meta-analysis. The search was limited to empirical studies employing quantitative methods, particularly experimental and quasi-experimental designs. Only articles written in English, with full-text availability and sufficient statistical data (e.g., means, standard deviations, and effect sizes), were included. The screening process involved several stages: initial title and abstract screening, full-text assessment, and duplicate removal using reference management software. Study selection was guided by the PICOS framework,

encompassing population, intervention, comparison, and contextual relevance, as outlined in [Table 1](#).
 outcome, and study design, to ensure methodological rigor

TABLE 1 | Characteristics of Included Studies

Author	Sample Size	Participant Characteristics	Intervention	Control	Score	Measurement Tools	Duration
Hidalgo-Camacho et al. (2019)	69 students (36 experimental, 33 control)	EFL beginner undergraduates in Ecuador	DI in reading comprehension using leveled materials	Traditional reading instruction	Improved post-test scores in the experimental group (quantitative not specified)	Pre-Post test and teacher observation	Not stated
Jefferson et al. (2017)	83 students (Grade 3, 6 classrooms)	Grade 3 students from three schools in Texas	Differentiated instruction, including core curriculum, evidence-based reading instruction, and repeated reading	Traditional core reading instruction	Significant improvement in comprehension and fluency; no exact scores reported	Reading comprehension, fluency tests, and high-stakes measures	One semester
Kotob & Abadi (2019)	20 students (10 low achievers, 10 high achievers)	Grade 4 students, mixed ability classroom, Tyre region, Lebanon	Differentiated instruction using flexible grouping, content/process/product differentiation	None (single group, experimental design only)	Marked improvement in low achievers' scores; stable scores for high achievers (quantitative scores not detailed)	Informal Reading Inventory (IRI), Checklist of Intelligence Strengths	10 days
Labordo Jr. (2024)	60 students (30 experimental, 30 control)	Grade 11 students, L2 students, Philippines	DI with a multiple intelligences strategy	Traditional instruction	Experimental group: middle-high school level; control: sixth-level comprehension	posttests (virtual platform assessment)	Not stated
Magableh & Abdullah (2020)	60 students (30 experimental, 30 control)	Grade 6 EFL students, two schools in Taibeh District, Jordan	Differentiated instruction: flexible grouping, tiered assignment, tiered instruction (content, process, product)	Traditional instruction (one-size-fits-all)	Statistical significance favoring the experimental group (specific score difference not detailed)	English grammar achievement test	12 weeks
Magableh & Abdullah (2021)	54 students (27 experimental, 27 control)	10th-grade students, mixed-ability classrooms	DI with homogeneous grouping and tiered assignments	One-size-fits-all using a textbook	Experimental group outperformed control (quantitative not specified)	Standardized reading assessments (SEM-R framework)	Not stated

Menson (2024)	26 (JEEP 4 Section 12, MSU-Maguindanao)	University students, 2nd semester AY 2013-2014, advanced elementary to low intermediate English proficiency	Differentiated Learning Instruction (DLI)	None (one-group pre-test post-test design)	Pre-test mean: 51.75; Post-test mean: 63.44; $z = 2.71, p < 0.05$	Pre-test and post-test achievement scores	Not stated
Sapan & Mede (2022)	26 students + 1 teachers	8th-grade secondary school students in Istanbul, Turkey	Differentiated Instruction (DI)	Not clearly described, comparative qualitative feedback only	Improved scores in FLA, FLM, and LA (quantitative, not detailed)	Pre-test and post-test speaking assessment	Not stated

Study Selection

The PRISMA protocol was followed across three phases: identification, screening, and eligibility determination. Duplicates were removed using Zotero, and studies were assessed according to PICOS criteria. To enhance clarity and minimize redundancy, subheadings such as Search Strategy, Data Extraction, and Risk of Bias were incorporated. Figure 1 illustrates the selection process. Records were initially screened by titles and abstracts, followed by a full-text review to assess eligibility. Inclusion criteria were guided by the PICOS framework, focusing on studies involving EFL students, DI as the intervention, traditional methods as the comparison, English language proficiency as the outcome, and quantitative study designs. Two independent reviewers conducted the screening, and disagreements were resolved through consultation with a third reviewer. Only empirical studies with complete statistical data were included, while theoretical papers, reviews, and studies lacking sufficient information were excluded.

Data Extraction

A standardized data extraction protocol was employed to ensure methodological consistency and enhance the accuracy of cross-study comparisons. Two independent reviewers systematically extracted key study details, including authorship, publication year, research design, sample characteristics, and geographical context, type of DI intervention, comparison group, outcomes related to English language proficiency, statistical values, and intervention duration. Any inconsistencies were resolved through consensus-building discussions, with a third reviewer consulted when agreement could not be reached. In cases where multiple reports of the same study existed, the most complete version was used to ensure reliability for meta-analysis.

Risk of Bias

The potential risk of bias was systematically assessed using the Cochrane Risk of Bias 2.0 tool to ensure methodological integrity. The framework evaluates five pivotal domains: randomization procedures, adherence to intended interventions, completeness of outcome data, accuracy of outcome measurement, and transparency in outcome

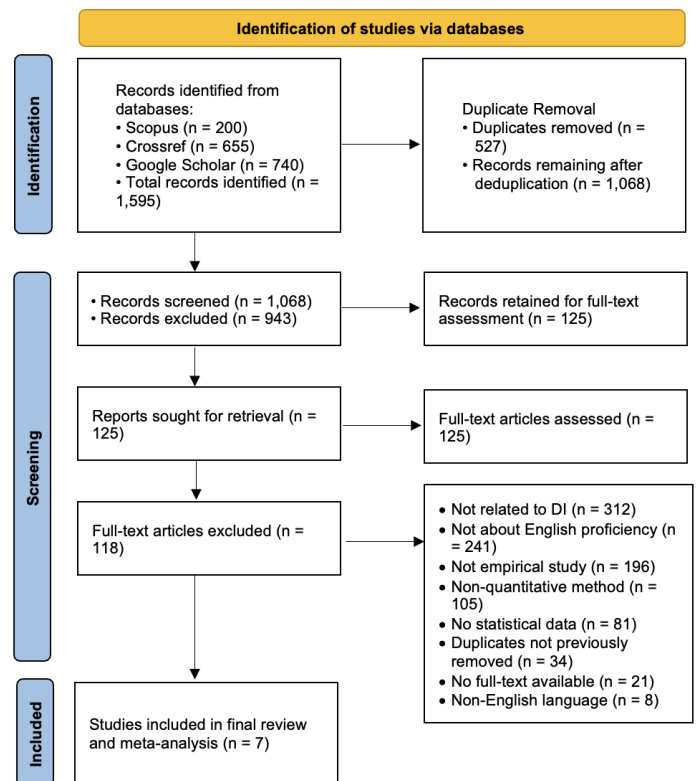


FIGURE 1 | PRISMA Flow Diagram of Article Selection

reporting. All assessments were documented in Review Manager (RevMan) version 5.4.1, which generated visual summaries such as traffic-light diagrams and summary graphs to facilitate interpretation and enhance transparency. The analysis indicated that while numerous studies demonstrated low risk in outcome measurement and selective reporting, significant methodological issues were identified, particularly in random sequence generation, allocation concealment, and blinding for participants and study personnel. These deficiencies pose potential threats to internal validity, suggesting that the results should be interpreted with caution. Consequently, the overall reliability of the meta-analytic findings depends on addressing these critical limitations in future research to strengthen methodological rigor and

validity.

Statistical analysis

Data analyses were conducted using RevMan version 5.4.1 to ensure methodological consistency and analytical accuracy. Effect sizes were calculated using the mean difference with a 95% confidence interval to compare outcomes between experimental and control groups. A random-effects model was applied in cases of significant heterogeneity, whereas a fixed-effects model was employed when heterogeneity was negligible. Heterogeneity was assessed using the Chi-square (χ^2) test and quantified with the I^2 statistic, where values of 25%, 50%, and 75% representing low, moderate, and high heterogeneity, respectively. Statistical significance was determined using the Z test, with results considered significant at p-value less than 0.05. Additionally, funnel plot analysis was conducted to visually examine and assess potential publication bias.

RESULTS AND DISCUSSION

The Search Results

The PRISMA flow diagram clearly depicts the systematic selection procedure implemented in this meta-analysis, which comprised ten structured steps. Initially, a comprehensive search identified 1,595 records from multiple databases: Scopus (n = 200), Crossref (n = 655), and Google Scholar (n = 740). After removing 527 duplicate, 1,068 records remained for screening based on titles and abstracts, resulting in the exclusion of 943 articles. The full texts of 125 articles were then assessed for eligibility, leading to the exclusion of 118 studies for various reasons; irrelevance to DI (n = 312), lack of focus on English language proficiency (n = 241), lack of an empirical methodology (n = 196), reliance on non-quantitative approaches (n = 105), insufficient statistical reporting (n = 81), overlooked duplicates (n = 34), unavailability of full-text versions (n = 21), and publications in languages other than English (n = 8). Ultimately, seven studies met all inclusion criteria and were incorporated into the final systematic review and meta-analysis. A detailed visual representation of this selection process is provided in Figure 1.

Characteristics of Included Studies

Table 1 presents a synthesized overview of seven empirical studies that implemented DI across various educational levels, addressing learners with diverse abilities and linguistic backgrounds. To ensure interpretability and transparency, this synthesis is supported by a PRISMA flowchart (Figure 1) outlining the study selection process, a forest plot (Figure 3) illustrating effect sizes across studies, and a funnel plot (Figure 4) highlighting potential publication bias. Each figure is described in detail within the narrative to ensure clarity, coherence, and adherence to meta-analytical reporting standards. Sample sizes varied from 20 students in Kotob and Abadi (2019) to 83 in Jefferson et al. (2017), with participants ranging Grade 4 learners to university-level students. The DI interventions included tiered assignments (Magableh & Abdullah, 2020; Magableh & Abdullah, 2020, 2021), flexible grouping (Kotob & Abadi, 2019), repeated reading (Jefferson

et al., 2017), and multiple intelligences strategies (Labordo, 2024). Menson (2024) reported significant learning gains among university students ($z = 2.71, p < 0.05$), while Jefferson et al. (2017) found improvements in comprehension and fluency. Measurement tools varied across the studies, and intervention durations ranged from 10 days to one semester. Overall, findings consistently showed that DI interventions outperformed traditional instructional methods in enhancing reading, grammar, fluency, and student engagement, confirming their effectiveness as summarized in Table 1.

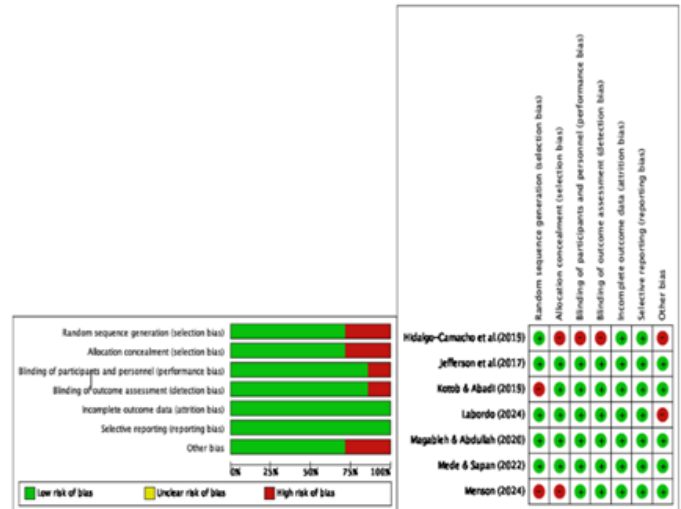


FIGURE 2 | Bias of Risk of 7 Studies

Outcome Measures

The outcome measures in this meta-analysis primarily focused on evaluating the effects of DI on various dimensions of English language proficiency. Across the included studies, assessments employed a diverse range of tools, including standardized reading assessments (e.g. SEM-R), grammar achievement tests, fluency tests, and pre- and post-test instruments tailored to specific skills such as reading comprehension, grammar, speaking, and overall language performance. For example, Jefferson et al. (2017) utilized comprehension and fluency tests; Kotob and Abadi (2019) applied Informal Reading Inventories and intelligence checklists; and Magableh and Abdullah (2020) employed grammar-specific assessments. Menson (2024) reported mean gains based on pre- and post-test evaluations. While these measurement tools varied in format and scope, they consistently demonstrated positive learning gains attributed to DI interventions, thereby validating the effectiveness of differentiated strategies across multiple linguistic domains and diverse students' profiles.

Study Quality

The assessment of study quality, as illustrated in Figure 2, shows notable variation in methodological rigor, particularly regarding the risk of bias. While most studies demonstrated low risk in areas such as outcome assessment, incomplete data, and selective reporting, high risk was most commonly observed in random sequence generation, allocation concealment, and blinding of participants and personnel. Specifically, over half of the studies did not provide sufficient detail regarding their randomization procedures, raising

concerns about potential selection bias. Similarly, inadequate allocation concealment and insufficient blinding increased the likelihood of performance and detection bias. Conversely, domains such as incomplete outcome data and selective reporting generally presented a low risk of bias. This overall pattern underscores the need for greater methodological rigor in future DI intervention studies, especially in implementing proper randomization and blinding procedures. For instance, studies such as [Hidalgo-Camacho et al. \(2019\)](#); [Kotob and Abadi \(2019\)](#); and [Menson \(2024\)](#) consistently showed high risk in these areas, indicating possible threats to internal validity. The accompanying bar chart further confirms that around 70% to 80% of the studies were affected by these critical biases. Additionally, a substantial proportion of "Other bias" suggests the presence of methodological issues not fully captured by standard categories. These findings emphasize the importance of interpreting the meta-analysis findings with caution, as weaknesses in randomization and blinding may have influenced the observed effect sizes. Recognizing these risks, as shown in [Figure 2](#), is essential for drawing valid and reliable conclusions.

Effects of interventions

Overall English Language Proficiency scores

The forest plot ([Figure 3](#)) shows that the pooled mean difference across seven studies was 2.92 with a 95% CI of [1.64-4.19], indicating that students who received DI outperformed those in control groups. This suggests that DI has a substantial positive impact on English proficiency. However, the high heterogeneity value ($I^2 = 87%$) indicates significant variability among studies, likely influenced by factors such as instructional duration and student demographics. The effect size is statistically significant, as evidenced by a Z-value of 4.49 and a p-value of < 0.00001 , confirming that the implementation of DI consistently yields higher outcomes than the control groups across studies. Among the included studies, [Jefferson et al. \(2017\)](#) and [Magableh and Abdullah \(2020\)](#) reported the largest effects (6.95 and 6.70, respectively), while the smallest effect was observed in [Hidalgo-Camacho et al. \(2019\)](#) with a mean

difference of 0.74. The variation in study outcomes is reflected by the substantial heterogeneity index ($I^2 = 87%$), suggesting that 87% of the observed variance in effect sizes is due to genuine differences across studies rather than random error. This is further supported by a significant Cochran's Q test ($Chi^2 = 45.74$, $df = 6$, $p < 0.00001$), confirming the presence of substantial between-study heterogeneity. Overall, these findings imply that while DI is generally effective, its magnitude of impact may vary depending on contextual factors such as sample characteristics, intervention design, and implementation fidelity. For a clearer understanding of the distribution and magnitude of effect sizes, see [Figure 3](#).

The assessment of bias and heterogeneity, as shown in [Figure 4](#), reveals notable methodological and statistical concerns. The asymmetry observed in the funnel plot suggests the presence of publication bias or small-study effects, possibly resulting from the selective non-reporting of studies with less favorable outcomes. Notably, the findings align with previous research emphasizing the positive impacts of DI on learning outcomes, particularly for students with diverse proficiency levels. DI enables educators to employ strategies such as differentiated grouping, instructional scaffolding, and stratified tasks tailored to learners' interests, preparedness, and cognitive profiles, thereby facilitating more effective management of mixed-ability classroom ([Maruf, 2023](#); [Nepal et al., 2021](#)). [Valiandes and Neophytou \(2018\)](#) highlight that these practices provide differentiated support and scaffolding based on students' readiness levels and leaning profiles. Similarly, [Puzio et al. \(2020\)](#) assert that differentiated literacy instruction is especially crucial in mixed-ability classrooms, given the growing diversity of educational contexts. Our findings also corroborate the work of [AM et al. \(2023\)](#), who found that differentiated strategies enhance academic performance and foster meaningful students engagement. Furthermore, the improvements in comprehension and fluency reported by [Smale-Jacobse et al. \(2019\)](#) resonate with our findings, emphasizing DI's role in fostering higher educational achievement and engagement through equitable learning opportunities ([Melese & Tinoca, 2019](#)).

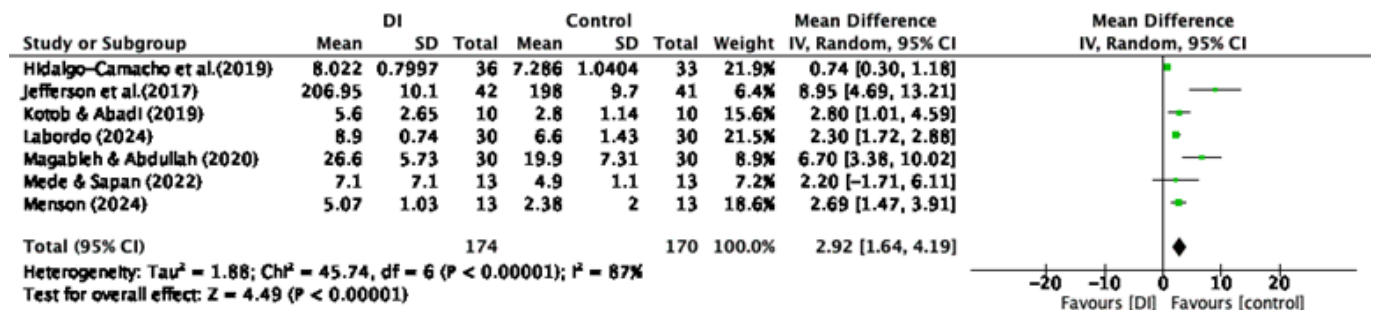


FIGURE 3 | Forest Plot Showing the Effect of Differentiated Instruction

However, our analysis highlights notable variances in effect sizes across studies, underscored by a high heterogeneity index ($I^2 = 87%$). This variability may stem from contextual factors such as differences in educational systems, student demographics, and the specific implementations of DI strategies. For instance, while [Puzio et al. \(2020\)](#) and [Mavidou and Kακαβά \(2019\)](#) reported larger

effect sizes, other studies observed much smaller effects, signaling the complex interplay of factors influencing DI's efficacy. This divergence supports the notion that while DI has strong potential to enhance learning, careful consideration of contextual and methodological factors is essential, echoing previous literature that similarly cautions against overgeneralizing findings from individual studies ([Mavidou](#)

& Kakana, 2019).

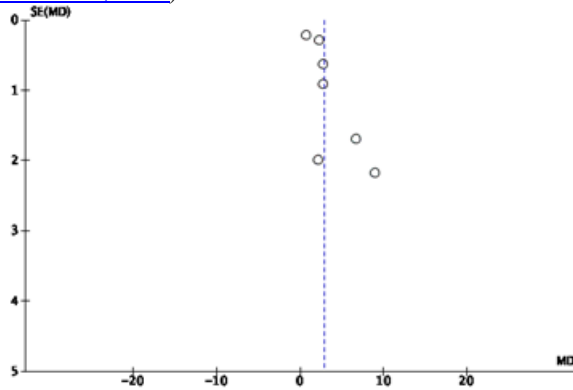


FIGURE 4 | Funnel Plot Indicating Publication Bias

The variation in the magnitude of learning gains across studies further indicates the importance of systematic and intentional implementation of DI. Studies employing well-structured experimental and control groups, as evidenced in [Jefferson et al. \(2017\)](#) and [Magableh and Abdullah \(2020\)](#) provided stronger empirical evidence than single-group pre-post-test designs ([Menson, 2024](#)). Also, students exposed to longer implementation periods, such as the 12-week intervention in [Magableh and Abdullah \(2020\)](#) showed greater improvements compared to the 10-days intervention in [Kotob and Abadi \(2019\)](#). These findings reinforce the idea that rigorous and deliberate alignment of teaching strategies with students' learning needs can become a transformative tool for addressing diverse student profiles.

Furthermore, the systematic review raises concern about methodological rigor, particularly risks of bias arising from inadequacies in sequence randomization and allocation concealment, as evidenced in studies such as [Kotob and Abadi \(2019\)](#) and [Uy \(2023\)](#). The short duration of some interventions and the absence of long-term follow-up assessments also raise questions about the sustainability of observed improvements. The high incidence of such biases calls for cautious interpretation of the findings, aligning with concerns raised in broader meta-analytic perspectives on educational interventions ([Suprayogi et al., 2024](#); [Uy, 2023](#)).

Moreover, four out of seven analyzed studies disproportionately emphasized certain domains of English proficiency, such as reading ([Hidalgo-Camacho et al., 2019](#); [Jefferson et al., 2017](#); [Kotob & Abadi, 2019](#); [Magableh & Abdullah, 2020](#)) and speaking ([Sapan & Mede, 2022](#)). Thus, while our results support the effectiveness of DI, the potential weaknesses in study designs warrant a cautious interpretation of these teaching methods, reinforcing calls for more rigorous future research ([Sahril et al., 2021](#); [Tricco et al., 2018](#)).

Although previous literature consistently supports DI's efficacy, this meta-analysis found that its effectiveness varies depending on implementation quality, instructional duration, and assessment methods. For example, studies using tiered assignments over extended periods produced greater learning gains. These findings highlight the need for strategic planning and teacher preparedness in applying DI effectively. Indeed, [Easa and Blonder \(2023\)](#) acknowledged that implementing DI practices is a complex and challenging task. Similarly, [Mavidou and Kakana \(2019\)](#) argue that curriculum

adjustments and the rigor of the implementation process significantly influence outcomes ([Mavidou & Kakana, 2019](#)). Our findings align with this perspective, as variations in duration, instructional strategies (e.g., tiered assignments vs. flexible grouping), and assessments types (standardized vs. high-stake tests) across studies indicate that DI's impact is nuanced and contingent upon proper execution ([AM et al., 2023](#); [Sahril et al., 2021](#)).

Moreover, our findings resonate with [Scott and Spencer's \(2006\)](#) assertion regarding the importance of intervention fidelity in achieving desired educational outcomes. Successful DI requires not only a strategic approach to instruction but also a strong commitment to quality implementation that aligns instructional strategies with the diverse needs of students ([Bullers et al., 2018](#)). The needs for intensive support for teachers, particularly in addressing challenges such as large class size, limited resources and access, time constraints in lesson preparation, has been highlighted as a reminder for policymakers ([Amoakwah & Donkoh, 2023](#); [Sellier & An, 2020](#); [Shareefa, 2021](#)). These findings corroborate [Menson's \(2024\)](#) earlier work, which emphasized the teachers must be adequately trained in DI practices to ensure effective application ([Uy, 2023](#)). Equally important is teachers' sense of efficacy, as it is positively associated with instructional quality, including the use of differentiated instruction and effective classroom management strategies ([Woodcock et al., 2022](#)).

When correlating our findings with previous research, it is notable that the push for DI is increasingly supported by educational policies advocating for more inclusive and adaptive teaching practices in mixed-ability classrooms. [Stone \(2018\)](#) highlights the necessity for teachers to move away from a "one-size-fits-all" model, mirroring our argument that DI enhances learning outcomes for diverse student populations ([Smale-Jacobse et al., 2019](#)). These findings underscore the pedagogical value of DI in addressing the linguistic and cognitive diversity commonly found in multilingual English classrooms, reinforcing its practical relevance for fostering inclusive and effective language learning environments.

In summary, our meta-analysis contributes to the growing body of literature substantiating the efficacy of DI in educational settings. The findings not only reinforce existing theoretical frameworks but also illuminate practical implications for educators. However, further research is critical to unpack the intrinsic variables influencing DI's effectiveness and to develop robust guidelines that can better support teachers in implementing these strategies across diverse contexts ([Partanen et al., 2019](#); [Varghese et al., 2021](#)). Sustained academic dialogue will be crucial in refining our understanding and approach to differentiated instruction in the pursuit of greater educational equity and excellence.

CONCLUSION

This meta-analysis substantiates the effectiveness of Differentiated Instruction in enhancing learners' grammar, reading comprehension, fluency, and overall English language proficiency. Compared to conventional pedagogical

approaches, DI consistently yields superior outcomes while fostering inclusivity in linguistically diverse classrooms. The evidence affirms DI's strategic role in advancing equitable and responsive language education.

However, the interpretation of these findings must be tempered by several methodological limitations, particularly inconsistencies in randomization procedures, allocation concealment, and blinding across the included studies. Furthermore, the substantial heterogeneity observed ($I^2 = 87\%$) underscores the need for caution when generalizing results beyond the reviewed contexts.

The implications of this synthesis advocate for the systematic integration of DI into language education policies and professional development programs. To reinforce the validity and applicability of future findings, subsequent research should adopt more rigorous experimental designs – particularly randomized controlled trials – and examine longitudinal effects, implementation fidelity, and contextual moderators such as teacher expertise and institutional support. Addressing these dimensions will not only refine theoretical understanding but also inform evidence-based instructional practices across diverse educational settings.

ACKNOWLEDGEMENTS

The authors wish to thank the institution for providing academic support throughout completing this article.

REFERENCES

- Ahmed, R. (2021). Peer review in academic writing: Different perspectives from instructors and students. *TESOL Journal*, 12(1). <https://doi.org/10.1002/tesj.537>
- AM, Muh. A., Hadi, S., Istiyono, E., & Retnawati, H. (2023). Does differentiated instruction affect learning outcome Systematic review and meta-analysis. *Journal of Pedagogical Research*. <https://doi.org/10.33902/JPR.202322021>
- Amoakwah, A., & Donkoh, S. (2023). basic school teachers' knowledge and use of differentiated instruction. *European Journal of Education Studies*, 10(10). <https://doi.org/10.46827/ejes.v10i10.5004>
- Bullers, K., Howard, A. M., Hanson, A., Kearns, W. D., Orriola, J. J., Polo, R. L., & Sakmar, K. A. (2018). It takes longer than you think: librarian time spent on systematic review tasks. *Journal of the Medical Library Association*, 106(2). <https://doi.org/10.5195/jmla.2018.323>
- Chien, C. W. (2015). Analysis of Taiwanese elementary school English teachers' perceptions of, designs of, and knowledge constructed about differentiated instruction in content. *Cogent Education*, 2(1), 16. <https://doi.org/10.1080/2331186X.2015.1111040>
- Easa, E., & Blonder, R. (2023). The development of an instrument for measuring teachers' and students' beliefs about differentiated instruction and teaching in heterogeneous chemistry classrooms. *Chemistry Teacher International*, 5(2), 125–141. <https://doi.org/10.1515/cti-2023-0005>

- Halil, N. I., Yawan, H., Hasanah, A. N., Syam, H., Andas, N. H., & Marhamah, M. (2024). A new program to foster inclusion: Unraveling language teachers' pedagogical practices to differentiated instruction. *International Journal of Language Education*, 8(2). <https://doi.org/10.26858/ijole.v8i2.64997>
- Hatmanto, E. D., & Rahmawati, F. (2023). Unleashing the potential: exploring attitudes and overcoming challenges in implementing differentiated instruction in the Philippines' English language classrooms. *E3S Web of Conferences*, 425, 02001. <https://doi.org/10.1051/e3sconf/202342502001>
- Hidalgo-Camacho, C. S., Elsa Hernandez Cherez, & Nelly Patricia Galora Moya. (2019). The effect of differentiated instruction on the development of reading comprehension skills in EFL beginner students. *Dominio de Las Ciencias*, 5(1), 125–144. <https://doi.org/10.23857/dc.v5i4.1045>
- Ismail, S. A. A., & Al Allaq, K. (2019). The nature of cooperative learning and differentiated instruction practices in English classes. *Sage Open*, 9(2). <https://doi.org/10.1177/2158244019856450>
- Jefferson, R. E., Grant, C. E., & Sander, J. B. (2017). Effects of tier I differentiation and reading intervention on reading fluency, comprehension, and high stakes measures. *Reading Psychology*, 38(1), 97–124. <https://doi.org/10.1080/02702711.2016.1235648>
- Kotob, M. M., & Ali Abadi, M. (2019). The Influence of Differentiated Instruction on Academic Achievement of Students in Mixed Ability Classrooms. *International Linguistics Research*, 2(2), p8. <https://doi.org/10.30560/ilr.v2n2p8>
- Labordo, D. C. (2024). Assessing the influence of differentiated Instruction on enhancing reading comprehension skills among learners: A quasi-experimental study. *The New English Teacher*, 18, 13–28.
- Lawrence-Brown, D. (2020). Differentiated Instruction and Inclusive Schooling. In Oxford Research Encyclopedia of Education. *Oxford University Press*. <https://doi.org/10.1093/acrefore/9780190264093.013>
- Magableh, I., & Abdullah, A. (2020). The Effect of Differentiated Instruction on EFL Learners: Teachers' Perspective. *International Journal of Academic Research in Business and Social Sciences*, 10(5). <https://doi.org/10.6007/IJARBS/v10-i5/7235>
- Magableh, I. S. I., & Abdullah, A. (2020). On the effectiveness of differentiated instruction in the enhancement of Jordanian students' overall achievement. *International Journal of Instruction*, 13(2), 533–548. <https://doi.org/10.29333/iji.2020.13237a>
- Magableh, I. S. I., & Abdullah, A. (2021). The impact of differentiated instruction on students' reading comprehension attainment in mixed-ability classrooms. *Interchange*, 52(2), 255–272. <https://doi.org/10.1007/s10780-021-09427-3>
- Mahali, S. C., & Sevigny, P. R. (2022). Multicultural Classrooms: Culturally Responsive Teaching Self-Efficacy among A Sample of Canadian Preservice

- Teachers. *Education and Urban Society*, 54(8), 946–968. <https://doi.org/10.1177/00131245211062526>
- Maruf, N. (2023). The interplay of teachers' beliefs, attitudes, and the implementation of differentiated instruction in Indonesian EFL contexts. *English Review: Journal of English Education*, 11(2), 357–364. <https://doi.org/10.25134/erjee.v11i2.7251>
- Mavidou, A., & Kakana, D. (2019). Teachers' experiences of a professional development program for differentiated instruction. *Creative Education*, 10(03), 555–569. <https://doi.org/10.4236/ce.2019.103040>
- Melese, S., & Tinoca, L. (2019). Instructors' knowledge, attitude and practice of differentiated instruction: The case of college of education and behavioral sciences, Bahir Dar University, Amhara region, Ethiopia. *Cogent Education*, 6(1). <https://doi.org/10.1080/2331186X.2019.1642294>
- Melka, Y. G., & Jatta, I. B. (2022a). Investigating the effects of differentiated instruction on high school students' grammar learning achievement in an Ethiopian EFL context, Addis Ababa University, Ethiopia. *East African Journal of Education Studies*, 5(3), 246–253. <https://doi.org/10.37284/eajes.5.1.636>
- Melka, Y. G., & Jatta, I. B. (2022b). Study the effects of differentiated instruction on English grammar teaching in Ethiopia. *Development of Education*, 5(3), 19–31. <https://doi.org/10.31483/r-103546>
- Menson, A. B. (2024). Differentiated learning instruction: Its effects on students' oral English fluency. *Journal of Data Acquisition and Processing*, 39(1), 1670. <https://doi.org/10.5281/zenodo.755230>
- Nassaji, H. (2017). Language instruction and language acquisition: A complex interplay. *Language Teaching Research*, 21(5), 543–545. <https://doi.org/10.1177/1362168817727203>
- Nepal, S., Walker, S., & Dillon-Wallace, J. (2021). How do Australian pre-service teachers understand differentiated instruction and associated concepts of inclusion and diversity? *International Journal of Inclusive Education*. <https://doi.org/10.1080/13603116.2021.1916111>
- Ojong, A. S. (2023). Unraveling the efficacy of differentiated instruction in enhancing second language acquisition: A comprehensive review and future directions. *International Journal of Linguistics, Literature and Translation*, 6(6), 75–82. <https://doi.org/10.32996/ijllt.2023.6.6.8>
- Onyishi, C. N., & Sefotho, M. M. (2020). Teachers' perspectives on the use of differentiated instruction in inclusive classrooms: Implication for teacher education. *International Journal of Higher Education*, 9(6), 136–150. <https://doi.org/10.5430/ijhe.v9n6p136>
- Ortega, D. P., Cabrera, J. M., & Benalcázar, J. V. (2018). Differentiating instruction in the language learning classroom: Theoretical considerations and practical applications. *Journal of Language Teaching and Research*, 9(6), 1220. <https://doi.org/10.17507/jltr.0906.11>
- Partanen, M., Siegel, L. S., & Giaschi, D. E. (2019). Longitudinal outcomes of an individualized and intensive reading intervention for third grade students. *Dyslexia*, 25(3), 227–245. <https://doi.org/10.1002/dys.1616>
- Puzio, K., Colby, G. T., & Algeo-Nichols, D. (2020). Differentiated literacy instruction: Boondoggle or best practice? *Review of Educational Research*, 90(4), 459–498. <https://doi.org/10.3102/0034654320933536>
- Sahril, S., N, A. S., & Nur, Muh. S. (2021). The impact of differentiated instruction on students' performance in critical reading. *Eralingua: Jurnal Pendidikan Bahasa Asing Dan Sastra*, 5(1), 275. <https://doi.org/10.26858/eralingua.v5i1.18937>
- Sapan, M., & Mede, E. (2022). The effects of differentiated instruction (DI) on achievement, motivation, and autonomy among English learners. *Iranian Journal of Language Teaching Research*, 10(1), 127–144. <https://doi.org/10.30466/ijltr.2022.121125>
- Sari, S. N. W. (2018). Providing explicit language instruction to assist students' language learning in EFL classrooms. *Indonesian Journal of EFL and Linguistics*, 3(1), 33. <https://doi.org/10.21462/ijefll.v3i1.55>
- Scott, W., & Spencer, F. (2006). Professional development for inclusive differentiated teaching practice. *Australian Journal of Learning Disabilities*, 11(1), 35–44. <https://doi.org/10.1080/19404150609546806>
- Sellier, N., & An, P. (2020). How peripheral interactive systems can support teachers with differentiated instruction: Using fireflies as a probe. *DIS 2020 - Proceedings of the 2020 ACM Designing Interactive Systems Conference*, 1117–1129. <https://doi.org/10.1145/3357236.3395497>
- Shareefa, M. (2021). Using differentiated instruction in multigrade classes: a case of a small school. *Asia Pacific Journal of Education*, 41(1), 167–181. <https://doi.org/10.1080/02188791.2020.1749559>
- Smale-Jacobse, A. E., Meijer, A., Helms-Lorenz, M., & Maulana, R. (2019). Differentiated instruction in secondary education: A systematic review of research evidence. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.02366>
- Spencer-Waterman, S. (2014). *Handbook on Differentiated Instruction for Middle & High Schools*. Routledge. <https://doi.org/10.4324/9781315856230>
- Stone, S. J. (2018). Differentiated Instruction: A Band-Aid Approach for a Flawed System. *International Journal of the Whole Child*, 3(1), 19–29.
- Suprayogi, M. N., Siregar, T. S., & Preston, M. (2024). The effectiveness of differentiated instruction, implementation in Indonesia higher education: A literature review (pp. 374–389). https://doi.org/10.2991/978-2-38476-108-1_37
- Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garrity, C., ... Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine*, 169(7), 467–473.

<https://doi.org/10.7326/M18-0850>

Uy, M. R. D. (2023). Differentiated instruction strategy and learners' performance. *International Journal of Research Publications*, 139(1).

<https://doi.org/10.47119/IJRP10013911220235802>

Valiandes, S., & Neophytou, L. (2018). Teachers' professional development for differentiated instruction in mixed-ability classrooms: investigating the impact of a development program on teachers' professional learning and on students' achievement. *Teacher Development*, 22(1), 123–138.

<https://doi.org/10.1080/13664530.2017.1338196>

Varghese, C., Bratsch-Hines, M., Aiken, H., & Vernon-Feagans, L. (2021). Elementary teachers' intervention fidelity in relation to reading and vocabulary outcomes for students at risk for reading-related disabilities. *Journal of Learning Disabilities*, 54(6), 484–496.

<https://doi.org/10.1177/0022219421999844>

Watts-Taffe, S., (Barbara) Laster, B. P., Broach, L., Marinak, B., McDonald Connor, C., & Walker-Dalhouse, D. (2012). Differentiated instruction: Making informed teacher decisions. *The Reading Teacher*, 66(4), 303–314.

<https://doi.org/10.1002/TRTR.01126>

Woodcock, S., Sharma, U., Subban, P., & Hitches, E. (2022). Teacher self-efficacy and inclusive education practices: Rethinking teachers' engagement with inclusive practices. *Teaching and Teacher Education*, 117, 103802.

<https://doi.org/10.1016/j.tate.2022.103802>

Zólyomi, A. (2022). Exploring Hungarian secondary school English teachers' beliefs about differentiated instruction. *Language Teaching Research*.

<https://doi.org/10.1177/13621688221114780>

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright ©2025 Riza Weganofa, Mujiono. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



From classroom to screen: Vocational English students' copywriting practices in creating tourism promotion videos

Dhini Aulia,* Ahmad Bukhori Muslim, Budi Hermawan

Universitas Pendidikan Indonesia, Indonesia

While existing research on copywriting primarily assesses the impact of promotional content on audiences, there is a notable lack of studies exploring how copywriting is taught and practiced within language education. This study addresses that gap by examining the strategies employed by English language students at a state polytechnic in Indonesia as they created tourism promotional videos. Utilizing a qualitative content analysis approach, the research investigates students' final reports to identify key patterns and gaps in their copywriting practices. The findings reveal that although students demonstrated creativity and effectively showcased local cultural and natural attractions through engaging language, their work often lacked depth in audience profiling, inclusion of practical travel information, and alignment between video duration and selected media platforms. These limitations reduced the communicative effectiveness of their content. The results suggest that, despite students' strong storytelling skills, they require more structured guidance on targeting, information inclusion, and media optimization in video copywriting. Integrating targeted copywriting instruction into English for Specific Purposes curricula, particularly in vocational contexts, has the potential to enhance students' readiness for professional promotional communication.

Keywords: copywriting strategies, English for Specific Purposes (ESP), tourism promotion, student-generated content

OPEN ACCESS

ISSN 2503 3492 (online)

*Correspondence:

Dhini Aulia

dhiniaulia.da@upi.edu

Received: 22th May 2025

Accepted: 14th August 2025

Published: 7th October 2025

Citation:

Aulia, D., Muslim, A.B., Hermawan, B. (2025). From classroom to screen: Vocational English students' copywriting practices in creating tourism promotion videos. *JEES (Journal of English Educators Society)*, 10(2). <https://doi.org/10.21070/jees.v10i2.1935>

INTRODUCTION

The field of copywriting has been examined through interdisciplinary research, with studies emerging in communication, psychology, business, and management (Karapenchev, 2020). The demand for copywriting continues to rise in the current era of social media, particularly in Asia and the Pacific region (Asian Development Bank & LinkedIn, 2022). This trend corresponds with the global shift in which companies are allocating larger portions of their budgets to advertising (Nielsen, 2024). In addition, a substantial proportion of both large and small businesses now use social media platforms to disseminate content with their overall brand messaging and promotional strategies (Mishra & Vijay, 2023).

The primary objective of copywriting is to motivate audiences to take desired actions, such as purchasing products, enrolling in programs, and reserving travel accommodations (Albrighton, 2013; Bly, 2005; Ogilvy, 2004; Sugarman, 2007). However, copywriting extends beyond persuasive writing; it involves the development of strategy, compelling ideas, and brand personality (Ogilvy, 2004) and is often described as both the art and science of crafting

persuasive written content (Noel & Wiebe, 2022). In today's business landscape, copywriters play a pivotal role in advertising and marketing, particularly in the domain of digital publications. Consequently, the demand for strong copywriting skills in the social media era is substantial.

A study by Chang and Kung (2023) found that copywriting strategies tailored to gender preference can enhance engagement with tourism products promoted on a Facebook fan page. Specifically, male audiences showed a preference for rational copywriting, whereas female audiences favored more emotional content. These findings align with the copywriting strategies framework proposed by Konstant (2011), Krawczyk (2022), and Ogilvy (2004), which emphasizes the significance of considering audience perspectives and emotional responses in copywriting. The framework synthesized in Table 1 integrates the key elements outlined by these three experts.

TABLE 1 | Copywriting strategies framework

Strategic Category	Description
Audience Awareness	Understanding the target audience and tailoring the message accordingly
Product Understanding	Developing in-depth knowledge of the product, including its benefits and features
Message Planning	Deciding what to communicate and how to structure the message
Crafting the Message	Writing with clarity, credibility, emotional appeal, and strategic focus
Reader Engagement	Grabbing attention, keeping interest, and calling to action
Testing and Revision	Reviewing and refining the copy to improve clarity and performance

Before composing a written piece, a copywriter must conduct extensive research on the target demographics likely to engage with the content. It is imperative to identify audience-specific factors such as culture, beliefs, and desires (Konstant, 2011). Numerous variables, including gender, age, marital or relationship status, children, opinions about the product, service, or industry, personality traits, brand preferences, and values, influence how a copywriter conveys messages (Krawczyk, 2022). Additionally, a copywriter must develop a detailed understanding of the specific product being promoted (Ogilvy, 1985).

As Albrighton (2013) asserts, different media require different communication strategies, and the choice of platform significantly shapes the language and style used. This study analyzes the copywriting language employed in promotional videos created by students and disseminated across platforms such as YouTube and Instagram. The effectiveness of video marketing in the tourism industry depends not only on creativity, but also on the adaptation of content to the unique characteristics of each platform. For instance, concise, vertical videos are better suited to Instagram Reels, whereas more in-depth, landscape-oriented

content works more effectively on YouTube.

Research has demonstrated the effectiveness of storytelling as a medium for conveying information, particularly in the context of travel and tourism. Well-crafted narratives in destination videos have been shown to enhance viewer engagement, communicate brand values, and cultivate long-term traveler loyalty by evoking emotional responses that leave a lasting impression and generate satisfaction. Keskin Yilmaz (2023) further points out that storytelling on official tourism channels, such as Türkiye's YouTube page, can influence traveler preferences and strengthen destination loyalty. Therefore, combining platform-specific adaptations with authentic storytelling has the potential to maximize the impact of digital promotional content in the tourism industry.

The linguistic nuances employed in copywriting also significantly influence audiences' willingness to act. Zulkifly and Firdaus (2014) discovered that in the case of a slimming product advertisement, audiences were more likely to respond when the headline was distinctive. Furthermore, linguistic choices can shape perceptions of authenticity, making it essential for messages to be clear and concise without resorting to exaggeration. Furthermore, Decrop (2007) indicates that while images are effective in capturing and stimulating interest, text remains the most powerful medium for conveying information. For products with a high degree of differentiation, he argues, visuals become less significant and less effective, particularly for transmitting detailed information.

Tourism destinations constitute a major sector of the world's most developed industries. Recent destination promotions have increasingly relied on social media platforms, particularly through promotional video, to persuade individuals to take action. A substantial body of research in psychology, marketing, business, and management demonstrates that short promotional videos in the tourism sector can serve as catalysts for travel-related behaviors (Keskin Yilmaz, 2023; Qiu et al., 2024; Wu & Ding, 2023; Zheng et al., 2023). Keskin Yilmaz (2023), for example, demonstrated that integrating storytelling into promotional videos on Türkiye's official YouTube channel can promote destinations, influence tourist preferences, and increase destination loyalty by improving visitor satisfaction.

As Ogilvy (2004) delineated, copywriting for tourist destinations is a distinct field of practice. He outlined several key principles: (1) destination advertising strongly influences the image of the target country; (2) it is essential to highlight the nation's distinctiveness; (3) advertisements should create a memorable image in the minds of readers; (4) audiences should not be underestimated, and adult language should be used instead of the clichés common in conventional travel advertising; (5) advertisements should justify travel costs by emphasizing cultural and status value; (6) the appeal of the destination should be highlighted; (7) texts should provide readers with specific, detailed instructions; and (8) care should be taken when addressing esoteric topics, since foreign tourists often prefer conventional and familiar representations.

In light of the aforementioned theories and research findings, it can be posited that integrating copywriting skills into the creation of promotional videos directly impacts the

effectiveness of the messages conveyed. The process of researching target audiences ([Chang & Kung, 2023](#); [Lapin, 2020](#); [Sofyan et al., 2023](#)), gaining a deep understanding of the product being promoted ([Krawczyk, 2022](#); [Ogilvy, 2004](#); [Zabukovec Baruca & Čivre, 2022](#)), and determining the appropriate medium and video duration ([Hou, 2017](#); [Petrova, 2025](#)) relies on copywriting strategies, which have been demonstrated to enhance audience engagement and prompt action ([Keskin Yilmaz, 2023](#); [Zulkifly & Firdaus, 2014](#)).

Despite a wealth practitioner-authored resources on copywriting, academic research on the subject remains limited, particularly within language learning classrooms ([Henson, 1996](#)). While numerous studies have explored the effects of promotional videos and copywriting on tourist perceptions and behaviors, relatively little attention has been paid how such videos are actually produced, especially in educational settings or by amateur creators such as students. Within the field of English for Specific Purposes (ESP), scholars have emphasized the importance of aligning language instruction with authentic, domain-specific communication tasks ([Hyland, 2002](#); [Basturkmen, 2010](#)). However, few studies have systematically analyzed the pedagogical approaches and practical processes involved in teaching or applying copywriting in ESP courses, particularly in vocational education.

To date, the existing literature has not sufficiently documented the strategies employed by students when producing promotional videos in ESP-based vocational classrooms, nor how these strategies reflect the realities of domain-specific communication. Therefore, this study seeks to fill this gap by investigating the strategies employed by vocational English language students in creating promotional videos intended for international tourists. Accordingly, the study aims to address the following research question:

“What pattern and features characterize the copywriting approaches used by vocational English language students in their self-produced promotional tourism videos?”

METHODS

The study aimed to provide a detailed analysis of the copywriting strategies used in promotional videos for tourism destinations produced by students enrolled in a Diploma 3 English program at an Indonesian state polytechnic in 2024. As part of the program requirements, students were expected to create English-based products, one of which was a video project, as their final assignment to complete their studies.

Approval was obtained from the head of the English Department at the state polytechnic to access students' final project materials. The researchers operated as independent academic investigators with no teaching responsibilities or institutional roles within the polytechnic, ensuring complete separation from any instructional or evaluative functions that might have influenced the participants. Since the study exclusively analyzed existing video projects and anonymized written reports, which posed minimal risk to participants, a waiver of individual student consent was requested and granted. Furthermore, the analysis revealed that several of these videos had already been publicly posted on YouTube by

their creators, further confirming the minimal risk and public availability of the data. Student identities were protected through the utilization of pseudonyms, and no identifying information was retained in the research datasets.

Moreover, this study employed a qualitative approach to examine the copywriting strategies included in the video-producing process. The content of students' final project reports was interpreted and analyzed utilizing qualitative content analysis to uncover the meaning embedded in the texts ([Drisko & Maschi, 2016](#); [Krippendorff, 2018](#); [Neuendorf, 2017](#); [Schreier, 2012](#)). The data collection followed [Mayring's \(2014\)](#) inductive category development, and the analysis applied [Kuckartz and Rädiker \(2023\)](#) evaluative qualitative content analysis. This approach allowed the researchers to assess how copywriting strategies were integrated into promotional videos created by students in a vocational English program. The process of evaluative qualitative analysis further enabled a nuanced understanding of the effectiveness and quality of the students' final projects.

In line with [Krippendorff \(2018\)](#), the following factors guided the choice of data sources: the ability of the videos and reports to provide comprehensive information on the research topic, their accessibility, and their focus on destination promotion. The dataset consisted of eight videos whose titles and reports explicitly referenced tourism destinations (see [Table 2](#)). All projects had been approved by the students' supervisor and the department examiners. The videos represented the final versions, which had also been authorized by the students for posting on YouTube or other social media platforms.

The study adapted a combined copywriting framework drawn from [Konstant \(2011\)](#), [Krawczyk \(2022\)](#), and [Ogilvy \(2004\)](#) (see [Table 1](#)) to analyze students' promotional video projects. In the context of ESP, Ogilvy's copywriting strategies - audience awareness, product understanding, and message planning - offer a pragmatic framework for teaching genre-specific promotional writing ([Hyland, 2022](#); [Basturkmen, 2010](#)). [Krawczyk \(2022\)](#) provided updated guidance on audience engagement tactics and multimedia adaptation, while [Konstant \(2022\)](#) offered foundational principles of persuasive language aligned with ESP learning outcomes in vocational communication. Despite the established presence of these models in professional copywriting literature and online resources, their integration into ESP teaching within media-production context remains limited. To address this gap, the analytical procedure was adapted to the research question and the nature of student-generated data. As illustrated in [Table 2](#), the final framework represents a novel, context-sensitive instrument for researchers and ESP instructors to evaluate and teach copywriting in video-based vocational English course.

In the context of qualitative content analysis, reports and videos were among the most common forms of data ([Drisko & Maschi, 2016](#); [Kuckartz & Rädiker, 2023](#)). This study employed participants' project reports as the unit of analysis, drawing on the framework and categories outlined in points 1, 2, and 3 of [Table 3](#). These categories were selected to address the research question concerning the copywriting strategies students used in creating promotional videos. NVivo 12 software was employed to facilitate the systematic

organization and coding of the data, thereby ensuring rigor and transparency in the analytical process.

TABLE 2 | The category of analysis

1	Target audiences	- gender - age - marital or relationship status - values: hobbies, religion, career, financial comfort, travel
2	Product	- Unique features of the country and destinations - Memorable imagery - Information on how to reach the destination - Estimated travel costs
3	Medium	- Published on social media platforms such as YouTube - Appropriate video duration

An expert perspective was deemed necessary to review the data and provide feedback on its reliability and validity. To this end, two experts in academic research were involved in triangulation to enhance the robustness of the study. A comprehensive description of the research design, research questions, data collection procedures, and data analysis was also provided to support their evaluation. The insights of these professionals were instrumental in strengthening the trustworthiness of the findings ([Lindheim, 2022](#)).

The final drafts of the students' reports and videos were compiled for analysis. The dataset consisted of eight videos and accompanying reports submitted by students to the study program administration. Prior to analysis, the data was coded to identify videos designed to promote tourist destinations to foreign or international tourists. [Table 3](#) delineates the exploration of tourist destinations in West Sumatra, Indonesia.

TABLE 3 | The purpose of the participants' video project

Project No	Project Title	Purpose of Project (in abstract)
1	The Promotional Video of Lereng Green View Tanjung Alai Solok	The purpose of this promotional video was to provide information about Lereng Green View, particularly for international visitors.
2	Promotional video of Georite Goa Kelelawar; A New Tourism Attraction in Padang	The purpose of this video was to promote the natural beauty of Goa Kelelawar to foreign and local tourists.
3	Discover Air Manis Hillside: Unveiling Villa, Cafe, and Resto in a Travel Feature Video	The goal was to increase tourist visits and enhance the destination's reputation as a worthwhile place to visit. The anticipated outcome is a boost in both local and international tourist engagement, driven by the video's reach and appeal across multiple platforms.
4	Making A Documentary video of Sungkai Green Park Ecotourism	This project was ntended to promote and provide more in-depth information about Sungkai Green Park as one of the tourist destinations in West Sumatra.
5	Making A Documentary video of Saribu Rumah Gadang I South Solok Regency	This project aimed to explore and showcase the charm of the Saribu Rumah Gadang area, a culturally rich tourist destination located in Nagari Koto Baru, Sungai Pagu District, South Solok Regency, West Sumatra.
6	Making A Promotional Video of Pagang Cafe at Bukit Gado-Gado Padang	The purpose of this video was to promote the Café and to introduce it to the wider community. Another goal was to provide information about one of the businesses in the city of Padang engaged in the culinary field.
7	Making A Profile Feature Video of Rumah Pakankamih	The purpose of this project was to introduce and promote the uniqueness and facilities provided to attract visitors.
8	Making a Documentary Video of Sarasah Tanggo	This video was created in response to the lack of information about this potential destination, especially for international visitors. In the digital era, information spreads quickly through advanced technology such as social media, where people worldwide can access a wide range of information from abroad.

RESULTS AND DISCUSSION

The analysis began with a review of the students' project reports. The objective of this phase was to meticulously examine each component of the report to identify the categories associated with each research question. This study identified several significant patterns in a sample of eight students' final project reports. Initially, the students produced a wide range of video types, including documentaries, features, and promotional videos. These types of videos were indicated in the titles of their final project reports.

Secondly, it is important to note that all reports adhered to a standardized format, encompassing sections or chapters such as the abstract, introduction, summary, procedure, results, discussion, conclusion, recommendations, and references. This consistent structure followed the guidelines provided in the study program's official handbook, which served as the primary reference for report preparation. The objectives of the students' videos projects were generally outlined in the abstract and introduction chapters. The objective of the phase of the study was to ensure that all selected data aligned with the categories relevant to the research objective (Krippendorff, 2018; Kuckartz & Rädiker, 2023). Investigating the purposes of the projects was essential because not all of the students' video projects were intended to promote a product or object (see Table 3). Several submitted projects were designed with the sole aim of providing information to the audience.

Thirdly, the study found that the procedure chapters detailed the steps involved in video production, covering the stages of pre-production, production, and post-production. In the pre-production phase, students engaged in activities such as generating ideas, conducting research, drafting frameworks, preparing equipment and budgeting. The production phase encompassed the shooting process and the conduction of interviews. The post-production phase involved a series of meticulous tasks, including selecting the best footage, creating the script, recording voice-overs, editing, adding subtitles, and mixing audio components. Moreover, in this chapter, students were generally required to consult at least two sources when developing their promotional videos. Some students referenced Fachruddin's (2017) framework for television video production, while others drew upon Wibowo's (2007) theoretical approach. Table 4 below illustrates the video types, the structure of the reports, and the procedures followed in creating the promotional videos.

The findings below indicated that not all videos created by students fell strictly under the category of promotional videos. However, it is important to note that other forms of videos, such as documentaries and features, also shared a similar objective: the promotion of tourist destinations. This observation is consistent with the prevailing perspective that documentary or feature videos can also be utilized for product promotion (Carter, 2021; Smith, 2023).

Moreover, the second finding showed the detailed articulation of project purposes, target audience, distribution platforms, and video duration in the students' abstracts and introductions. This indicated that the students prepared their strategies before creating the videos. This approach aligns

with Fachruddin (2017), Wibowo (2001), and Mowat (2021), who emphasized that effective video production and marketing hinge on meticulous pre-production planning, encompassing clear objectives, audience analysis, and strategic distribution. In addition, research by Roy et al. (2021) demonstrated the application of pre-production planning in an educational setting, where students designed promotional videos as part of a project-based language learning curriculum. Their study highlighted how activities such as storyline development, screenplay writing, and storyboarding during the pre-production phase could effectively promote a product, in this case, intelligent toilets, while simultaneously enhancing students' professional communication skills. Segura (2023) also discussed the importance of pre-production planning with a promotional focus in the digital age, emphasizing that understanding the target audience, defining the video's purpose, and planning content repurposing strategies are crucial steps for maximizing a video's reach and impact.

TABLE 4 | Video type

No	Project Title	Video Type
1	The Promotional Video of Lereng Green View Tanjung Alai Solok	Promotional
2	Promotional Video of Geosite Goa Kelelawar; A New Tourism Attraction in Padang	Promotional
3	Discover Air Manis Hillside: Unveiling Villa, Cafe, and Resto in a Travel Feature Video	Feature
4	Making A Documentary Video of Sungkai Green Park Ecotourism	Documentary
5	Making A Documentary Video of Saribu Rumah Gadang in South Solok Regency	Documentary
6	Making A Promotional Video of Pagang Cafe at Bukit Gado-Gado Padang	Promotional
7	Making A Profile Feature Video of Rumah Pakankamih	Feature
8	Making A Documentary Video of Sarasah Tanggo	Documentary

The third finding illustrated the procedures used in creating promotional videos. The students followed the steps outlined by Fachrudin (2017) and Wibowo (2001), encompassing the phases of pre-production, production, and post-production. In the early stage of pre-production, the majority of reports described five steps: generating ideas, conducting research, drafting frameworks, preparing equipment, and budgeting. In line with the research objectives, this study examined the extent to which the target audience, product description, and video distribution were delineated in the pre-production phase of the students' final project reports. As noted by Konstant (2011), Krawczyk (2022), and Ogilvy (2004), and further supported by Mowat

(2021) and Segura (2023), a comprehensive understanding of the target audience, a clear and detailed presentation of the product, and a carefully formulated distribution strategy are indispensable components of effective video marketing campaigns. In the following subsections, an exposition of the three aforementioned components of copywriting strategy is provided.

Target Audience Description

A comprehensive review of the data indicated that the majority of reports identified foreign or international tourists as the primary target audience, as evidenced by the use of English voiceovers in the videos. Several words and phrases were frequently employed by students to describe the target audience. The most commonly used terms included "audience," "tourists," "viewers," and "visitors." Furthermore, the phrases most often appearing were "local and international tourists," "wider audience," "foreign tourists," and "international tourists." As demonstrated in Figure 1, generated by NVivo 12, the following data represents the word frequency results describing the target audience in students' project reports.

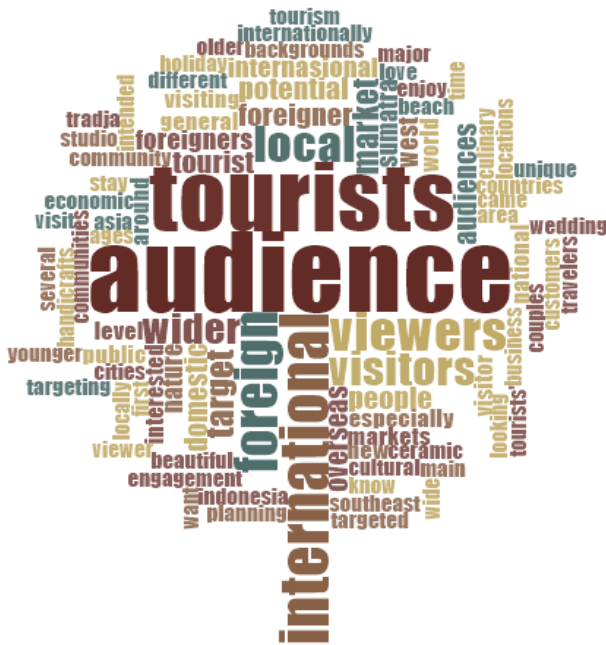


FIGURE 1 | Word frequency query of the target audience

An analysis of the target audience description revealed that some students provided only minimal additional detail. One report identified the target audience as "tourists planning to visit a cultural tourism area," while other states that it was "tourists who love the beach." Furthermore, one report included numerical data concerning tourists who had arrived at the destination, along with the types of tourists targeted, such as "holiday tourists, culinary tourists, nature tourists, wedding couples, business tourists, and especially foreign tourists visiting West Sumatra who are interested in nature."

The data suggested that students who created promotional videos to promote either tourist destinations or unique products were aware of the intended recipients of their videos. It was evident that the videos were intended for an international audience, as they were accompanied by English

voiceovers. However, there was a lack of details regarding the target audience's gender, age, and occupation or economic level. Consequently, this finding did not support the recommendations of several studies that emphasized the importance of providing detailed target audience profiles. For instance, Chang and Kung (2023) determined that gender differences among target audiences significantly affected engagement levels. Additionally, Rodrigo-Martín et al. (2025) emphasized the importance of identifying the target audiences' age, finding that adolescents were a crucial demographic whose engagement with promotional products depended on the extent to which the content encouraged interaction, empathy, and enjoyment in advertising communications.

Nevertheless, more precise characterization of the audience in the findings, such as "tourists planning to visit a cultural tourism area" and "tourists who love the beach", proved more pivotal in effectively conveying a compelling message. The writing style and word choices employed in the video transcripts aligned with the specific information needs of the target audience. This approach aligns with the findings of Sofyan et al. (2023), who noted that comprehensive information about the target audience significantly influenced writing style. Furthermore, in the domain of copywriting, the efficacy of the content was significantly shaped by word choice (Lapin, 2020).

Product Descriptions

The second category analyzed to address the research questions was the presentation of the product. This category involved examining several specific points that described detailed information about either tourist destinations or unique products in the data. A comprehensive analysis of eight reports on tourist destinations was conducted, encompassing various subcategories. These subcategories followed Ogilvy's framework, which includes the country's image, destination image, cost of the journey, and information about how to reach the destination.

The first point analyzed in relation to tourist destination descriptions was the country's image, in this case, Indonesia. The analysis revealed that, while not all reports explicitly mentioned the country's image, a significant proportion, six out of eight, did so by highlighting its natural beauty. The phrase "natural beauty" frequently appeared in the introduction chapter. Furthermore, some reports described Indonesia's distinctive image, emphasizing its flora, fauna, forests, beaches, lakes, waterfalls, and cultural potential. Some video scripts included in the reports conveyed Indonesia's image.

West Sumatra emerged as the second most prevalent focus in discussions of the nation's image. The reports provided detailed descriptions of the region's natural beauty and numerous tourist attractions. Additionally, the concept of cultural tourism was frequently referenced in the reports that discussed the country's image, with the term most commonly found in the introduction chapters.

The subsequent aspect of destination description pertained to the uniqueness of the destinations. This category was predominantly discussed in the introduction section and reflected in the video scripts presented in the result section. A

comprehensive analysis of eight final reports revealed that all of them highlighted destination uniqueness by underscoring distinctive features such as architectural elements, facilities, cuisine, and overall atmosphere. Content analysis further indicated that uniqueness emerged as the most prevalent theme, surpassing other categories. Furthermore, the type of video influenced the extent of detail provided. For instance, a documentary-style video tended to deliver a more comprehensive account of a destination's historical and cultural nuances than a purely promotional video.



FIGURE 2 | The country's image described the natural and cultural dimensions

The terms "natural," "beauty," "culture," and "cultural" were predominantly employed to characterize Indonesia, the focal nation in this study. According to the reports, destinations were consistently described in detail in the video scripts. This finding aligns with [Ogilvy \(2004\)](#) principles of copywriting for promoting tourist destinations. The distinctiveness of destinations in this study, emphasizing cultural and natural dimensions as illustrated in [Figure 2](#), also aligns with the findings of [Shani et al. \(2010\)](#) and [Zabukovec Baruca and Čivre \(2022\)](#), who reported that attributes such as natural resources, cultural heritage, and culinary experiences significantly motivate tourists throughout the year.

However, the analysis revealed that essential information, such as the cost of the journey, was not addressed in the Introduction or Results sections of the reports, nor in other sections. Similarly, practical travel information was described only superficially. Specific details regarding the destination's geographical context, such as its location in West Sumatra, operating hours, and distance from Padang, the provincial capital, were mentioned only briefly.

Consequently, the data did not provide sufficient details on the cost of travel. Prior research has established that accessibility and affordability are pivotal in shaping tourists' decisions when selecting destinations ([Hou et al., 2016](#)). Price, in particular, is a decisive factor, as tourists often evaluate affordability in relation to the perceived value of the

experience ([Levyda et al., 2022](#)). Therefore, incorporating cost-related information into promotional materials is strongly recommended to better support prospective visitors in evaluating potential destinations ([Jiang et al., 2022](#)).

Furthermore, information on practical aspects of the journey was not explicitly presented in the video scripts or in other sections of the reports. These findings stand in contradict to established theoretical standards for content in tourist promotional videos. Prior research indicates that including transportation options, such as air, land, or water travel, enhances potential visitors' understanding of the ease and convenience of reaching the destination. For instance, studies on promotional videos from China ([Hou et al., 2016](#)) demonstrate that most videos include details on both land and air transportation, underscoring their impact on tourists' decision-making. Such information helps address logistical concerns and ultimately makes destinations more appealing.

Media to Publish and Video Duration

Identifying the optimal medium for disseminating copywriting products is essential to ensure that the intended audience engages with the promotional video at the right time, thereby maximizing its impact and achieving marketing objectives. The importance of media selection is supported by research highlighting the need to align advertising strategies with audience preferences and behaviors ([Deu & Baharudin, 2024](#)).

The duration of the students' video projects ranges from three to seven minutes, depending on the nature of the content. Documentary videos were generally longer than promotional or profile videos. Regarding dissemination, four projects in this study explicitly identified Instagram and YouTube as their chosen publication platforms. Two reports indicated a preference for digital or social media more broadly, while the remaining reports did not specify the intended venue for dissemination. As summarized in [Table 5](#), the students' projects are outlined with details on the selected media platforms and video duration.

Additionally, the data provided insights into the duration of the videos, which ranged from three and a half to ten minutes. The findings indicated that the video length was not associated with the choice of publication platform or the type of video. A notable disparity in duration was observed cross the project videos. For instance, a comparative analysis of the three videos published on YouTube revealed that the Sungkai Green Park Ecotourism video had a considerably longer duration than the other two, Sarasah Tanggo and Geosite Goa Kelelawar, each of which lasted approximately four minutes. By contrast, promotional videos generally ranged between three and four minutes, while documentary videos varied more substantially in length, with examples lasting ten, seven, and four minutes.

This study found that most reports indicated publication on major digital platforms such as YouTube, Instagram, and other social media networks. While all reports specified video duration, none addressed the relationship between video length and the chosen platform. This suggests that students lacked awareness of distinct audience behaviors influencing optimal video length on different platforms ([Petrova, 2025](#)).

TABLE 5 | Media to publish and video duration

Project Title	Media to Publish	Video duration
Making a Documentary Video of <i>Saribu Rumah Gadang</i> in South Solok	Not mention	10 minutes and 7 seconds
Making A Promotional Video of Pagang Cafe at Bukit Gado-Gado Padang	Instagram	3 minutes and 39 seconds
Promotional Video of Geosite Goa Kelelawar; A New Tourism Attraction in Padang	YouTube	4 minutes and 23 seconds
The Promotional Video of <i>Lereng Green View</i> Tanjung Alai Solok	Not mention	4 minutes and 15 seconds
Discover Air Manis Hillside: Unveiling Villa, Cafe, and Resto in a Travel Feature Video	Social media	7 minutes and 2 seconds
Making A Profile Feature Video of Rumah Pakan Kamih	Digital media	4 minutes and 44 seconds
Making a Documentary Video of Sungkai Green Park Ecotourism	YouTube	7 minutes and 15 seconds
Making a Documentary Video of Sarasah Tanggo	YouTube	4 minutes and 37 seconds

However, the analysis of video content in this study reveals that students' duration choices were consistent with destination marketing research by [Hou et al. \(2016\)](#). They found that many official tourism videos in China, especially those produced by governments agencies or destination management organizations, employed longer formats. These videos offered comprehensive introductions to local culture, festivals, transportation, accommodations, shopping, and unique experiences, often using narrative techniques such as documentary style, whole-to-part sequencing, day-to-night transitions. The effectiveness of such videos lay in their strong internal logic and extensive coverage, which ranged from several to over 20 minutes. This breadth proved valuable for audiences in the planning stage, who sought detailed and holistic destination insights. Similarly, the videos analyzed in the present study emphasized Indonesian cultural features, architecture, traditional cuisine, and natural landscapes, content that required extended duration for effective storytelling.

These findings align with those of ESP digital project studies conducted in other contexts. [Apriyanti et al. \(2021\)](#) found that Indonesian vocational students struggled with audience awareness in ESP video production, while [Klimova & Kacatl \(2015\)](#) reported that Czech students exhibited strong technical skills but weaker strategic communication. In contrast, the current study revealed stronger integration of cultural content but limited optimization of media platforms. In a similar vein, [Girón-García & Fortanet-Gómez \(2023\)](#) found that Spanish students effectively utilized multimodal ESP resources with sophisticated audience targeting. By comparison, the Indonesian students in this study produced rich content but demonstrated weaker alignment between platform selection and audience alignment.

Moreover, the findings suggest that incorporating comprehensive copywriting strategies into the development of tourism promotional videos can significantly enhance the effectiveness of destination marketing, particularly when such strategies are guided by a thorough understanding of the target audience and the distinctive attributes of the

destination. The findings indicate that when students employed targeted messaging, vivid product descriptions, and appropriate media selection, their promotional content become more persuasive and engaging for international audiences. As a result, students' copywriting choices directly contributed to language learning gains. Extended video duration, for example, required the use of complex descriptive language, comparative structure, and persuasive devices-linguistic features that align closely with ESP competencies.

Building on these findings, the creation of promotional videos by English language students in an ESP context reveals the pedagogical value of integrating authentic, task-based activities that reflect real-world professional demands. In particular, embedding copywriting for tourism promotion into the ESP curriculum not only enhances students' linguistic competence but also fosters strategic thinking, creativity, and audience awareness, core competencies in vocational communication. This aligns with the overarching goal of ESP; to equip learners with language and discourse practices relevant to their future professions ([Belcher, 2009](#); [Hyland, 2022](#)). By engaging students in the production of tourism videos, educators can provide experiential learning opportunities that connect language instruction with marketing, media literacy, and intercultural communication. Furthermore, such tasks create fertile ground for curriculum developers to design modules that incorporate message targeting, persuasive language use, multimedia scripting, and digital content adaptation across platforms. Embedding these practices in ESP instruction encourages students to move beyond surface-level language use and apply critical communication strategies, thereby preparing them for workplace scenarios where English functions as both a communicative and persuasive tool.

CONCLUSION

This study explored the copywriting strategies employed by Diploma 3 English language students at an Indonesian state

polytechnic in producing promotional videos for tourist destinations in West Sumatra. The findings revealed that, while students demonstrated creativity and effectively highlighted the distinctive features of Indonesian destinations through engaging language, their copywriting often lacked depth in audience profiling, practical travel information, and alignment between video duration and chosen media platforms. These aspects are essential to ensuring that promotional messages are not only appealing but also relevant and actionable for the intended audience. The results underscore the need to integrate comprehensive audience analysis, practical content, and platform-specific strategies into language education curricula to better prepare students for real-world digital communication tasks in tourism promotion. Addressing these areas in future educational and training initiatives could enhance both the impact and effectiveness of student-produced promotional materials within the tourism sector.

However, while this study provides valuable insights into the copywriting strategies employed by English language students in producing promotional videos, several limitations must be acknowledged. First, the research was confined to a single institution, which restricts the generalizability of the findings. Second, the study did not investigate students' prior knowledge, experiences, or curriculum exposure that may have shaped their approach to copywriting and video production. Notably, this study addresses a significant gap in the existing literature: whereas most previous research has focused on the effects of copywriting products on audience attitudes and behaviors, the processes through which these products are conceptualized, researched, and crafted remains largely unexplored. Understanding the stages of research, idea generation, and script development is crucial, as effective copywriting depends on systematic research and creative processes to produce persuasive and relevant content. Therefore, future research should not only examine the outcomes of copywriting but also explore the educational and experiential factors that influence how copywriting products are produced, offering a more comprehensive view of copywriting in both academic and professional contexts.

ACKNOWLEDGEMENTS

The authors would like to express their sincere gratitude to all participants and colleagues who contributed to this research. This study was supported by the Indonesia Education Scholarship (Beasiswa Pendidikan Indonesia). The funding agency had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

REFERENCES

- Albrighton, T. (2013). *The ABC of copywriting* (2nd ed.). ABC Business Communications Ltd. <https://www.abccopywriting.com/wp-content/uploads/2013/04/The-ABC-of-Copywriting.pdf?utm>
- Apriyanti, D.-, Syarif, H., & Ramadhan, S. (2021). Video Feature Making in ESP- Based Public Speaking Class: A Student-Centred Learning in Vocational Higher Education Context. *International Journal of Language Education*, 469–476. <https://doi.org/10.26858/ijole.v5i1.15419>
- Asian Development Bank, & LinkedIn. (2022). Digital jobs and digital skills (Join Report ADB and LinkedIn SPR220348). <https://doi.org/10.22617/SPR220348>
- Basturkmen, H. (2010). *Developing courses in English for specific purposes*. Palgrave Macmillan. <https://doi.org/10.1057/9780230290518>
- Belcher, D. (2009). *What ESP is and can be: An introduction*. University of Michigan Press. <http://www.press.umich.edu/titleDetailDesc.do?id=770237>
- Bly, R. W. (2005). *The copywriter's handbook* (3rd ed.). Owl Books. ISBN: 978-0-8050-7804-6. <https://www.amazon.com/Copywriters-Handbook-Step-Step-Writing/dp/0805078045>
- Carter, A. (2021). Capturing authenticity: An exploration in brand-sponsored documentary video storytelling to define brand values and further social impact (Master's thesis, University of North Carolina at Chapel Hill). <https://doi.org/10.17615/339n-n465>
- Chang, H.-L., & Kung, Y.-H. (2023). Copywriting strategies in tourism Facebook fan page marketing: Investigating the role of gender differences. *Management Review (Taiwan)*, 42(2). [https://doi.org/10.6656/MR.202304_42\(2\).ENG001](https://doi.org/10.6656/MR.202304_42(2).ENG001)
- Decrop, A. (2007). The influence of message format on the effectiveness of print advertisements for tourism destinations. *International Journal of Advertising*, 26(4), 505–525. <https://doi.org/10.1080/02650487.2007.11073030>
- Deu, I., & Bakhtiar, M. B. (2024, April). Purchase Intention and Marketing Media: A Comparison Between Photos and Videos. In *CoMBInES-Conference on Management, Business, Innovation, Education and Social Sciences* (Vol. 4, No. 1, pp. 283-290). https://scholar.google.com/scholar?hl=id&as_sdt=0%2C5&q=Deu%2C+I.%2C+%26+Baharudin%2C+M.+%282024%29.+Purchase+intention+and+marketing+media%3A+A+comparison+between+photos+and+videos.+%5BConference+Paper%5D.&btnG=
- Drisko, J. W., & Maschi, T. (2016). *Content analysis*. Oxford University Press. <http://lcn.loc.gov/2015040785>
- Fachruddin, A. (2017). *Dasar-dasar produksi televisi: Produksi berita, feature, laporan investigasi, dokumenter dan teknik editing*. Kencana Prenadamedia. https://books.google.com/books/about/Dasar_dasar_Produksi_Televisi_Produksi_B.html?id=0KRPDwAAQBAJ
- Girón-García, C., & Fortanet-Gómez, I. (2023). Science dissemination videos as multimodal supporting resources for ESP teaching in higher education. *English for Specific Purposes*, 70, 164–176. <https://doi.org/10.1016/j.esp.2022.12.005>

- Henson, L. (1994). A preliminary rhetoric of technical copywriting. *Technical Communication*, 41(3), 447–455. <https://www.jstor.org/stable/43088472>
- Hou, Z. (2017). Comparison study of advertising videos of tourism city image—Based on the content analysis method. *EURASIA Journal of Mathematics, Science and Technology Education*, 13(10). <https://doi.org/10.12973/ejmste/78521>
- Hou, Z., Joppe, M., Choi, C., & Lin, Z. (2016). Exploratory study of promotional videos in the 10 major tourist destinations in China: A content analysis. *Travel and Tourism Research Association: Advancing Tourism Research Globally*, 45. <https://scholarworks.umass.edu/ttra/2011/Visual/45>
- Hyland, K. (2002). Specificity revisited: How far should we go now? *English for Specific Purposes*, 21(4), 385–395. [https://doi.org/10.1016/S0889-4906\(01\)00028-X](https://doi.org/10.1016/S0889-4906(01)00028-X)
- Hyland, K. (2022). English for specific purposes: What is it and where is it taking us? *ESP Today*, 10(2), 202–220. <https://doi.org/10.18485/esptoday.2022.10.2.1>
- Jiang, J., Hong, Y., Li, W., & Li, D. (2022). A study on the impact of official promotion short videos on tourists' destination decision-making in the post-epidemic era. *Frontiers in Psychology*, 13, 1015869. <https://doi.org/10.3389/fpsyg.2022.1015869>
- Karapenchev, Y. (2020). Defining copywriting within the framework of integrated marketing communications. In *Communication Management: Theory and Practice in the 21st Century*. https://research.uni-sofia.bg/bitstream/10506/2493/1/FJMC_Sbornik_Ko_nferentsia_New.pdf
- Keskin Yilmaz, Y. (2023). Storytelling in destinations' promotional videos: The case of Go Türkiye official YouTube page. *İNİF E-Dergi*. <https://doi.org/10.47107/inifedergi.1333502>
- Klimova, B. F., & Kacatl, J. (2015). Hybrid Learning and its Current Role in the Teaching of Foreign Languages. *Procedia - Social and Behavioral Sciences*, 182, 477–481. <https://doi.org/10.1016/j.sbspro.2015.04.830>
- Konstant, T. (2011). Ten rules of copywriting. Hodder Education. <http://www.hoddereducation.co.uk>
- Krawczyk, N. (2022). Copywriting strategies: A no-nonsense guide to writing persuasive copy for your business. Rockridge Press. <https://play.google.com/store/books/details?id=32IXEQAAQBAJ>
- Krippendorff, K. (2018). *Content analysis: An introduction to its methodology (4th ed.)*. SAGE Publications. <https://doi.org/10.4135/9781071878781>
- Kuckartz, U., & Rädiker, S. (2023). *Qualitative content analysis: Methods, practice and software (2nd ed.)*. SAGE Publications. <https://uk.sagepub.com/en-gb/asi/qualitative-content-analysis/book282907>
- Lapin, J. (2020, April 3). 3-step guide to creative copywriting. <https://www.archeredu.com/hemj/guide-to-creative-copywriting/>
- Levyda, L., Eddyono, F., & Meyrza, C. T. (2022). The effect of facility promotion and prices on the decision to visit DUFAn Ancol. *Jurnal Ekonomi*, 11(01). <http://ejournal.seaninstitute.or.id/index.php/Ekonomi>
- Lindheim, T. (2022). Participant validation: A strategy to strengthen the trustworthiness of your study and address ethical concerns. In G. Espedal, B. Jelstad Løvaas, S. Sirris, & A. Wæraas (Eds.), *Researching values* (pp. 225–239). Springer. https://doi.org/10.1007/978-3-030-90769-3_13
- Mayring, P. (2014). *Qualitative content analysis: Theoretical foundation, basic procedures and software solution*. Klagenfurt. <https://nbn-resolving.org/urn:nbn:de:0168-ssao-395173>
- Mishra, A., & Vijay, T. S. (2023). *Integrated advertising, promotion, and marketing: Communicating in a digital world (1st ed.)*. Routledge India. <https://doi.org/10.4324/9781003458593>
- Mowat, J. (2021). *Video marketing: Create engaging video campaigns to drive brand growth and sales (2nd ed.)*. Kogan Page. <https://www.koganpage.com/product/video-marketing-9781398601147>
- Neuendorf, K. A. (2017). *The content analysis guidebook (2nd ed.)*. SAGE. <http://lccn.loc.gov/2015044657>
- Nielsen. (2024). The advertising media spend report (Report H1 2024). The Nielsen Company (US). <https://www.nielsen.com/insights/2024/advertising-media-spend-report-h1-2024/>
- Noel, J., & Wiebe, J. (2022). What is copywriting? Why it's critical to your bottom line. Copywriting Articles. <https://copyhackers.com/2022/03/what-is-copywriting/>
- Ogilvy, D. (2004). *Confessions of an advertising man (Reprint ed.)*. Southbank. https://openlibrary.org/books/OL8774119M/Confessions_of_an_Advertising_Man
- Petrova, N. (2025, February 21). How to use video for tourism marketing. Video Creation. <https://firework.com/blog/how-to-use-video-for-tourism-marketing>
- Qiu, L., Li, X., & Choi, S. (2024). Exploring the influence of short video platforms on tourist attitudes and travel intention: A social–technical perspective. *Journal of Destination Marketing & Management*, 31, 100826. <https://doi.org/10.1016/j.jdmm.2023.100826>
- Rodrigo-Martín, I., Rodrigo-Martín, L., & Muñoz-Sastre, D. (2025). Publicidad y adolescencia. Estudio etnográfico de una comunidad virtual. *Index. Comunicación*, 15(1), 31–52. <https://doi.org/10.62008/ixc/15/01Public>
- Roy, D., Kosai, H., & Takai, K. (2021). The pre-production plans on designing a promotional video for project-based language learning: A case study with smart toilets. *SHS Web of Conferences*, 102, 02009. <https://doi.org/10.1051/shsconf/202110202009>
- Schreier, M. (2012). *Qualitative content analysis in practice*. SAGE Publications. <https://doi.org/10.4135/9781529682571>
- Segura, S. L. (2023, September 27). Video promotion: Preproduction strategies & marketing tips. AZoNetwork. <https://www.azonetwork.com/marketing-science/blog/video-promotion>

- Shani, A., Chen, P., Wang, Y., & Hua, N. (2010). Testing the impact of a promotional video on destination image change: Application of China as a tourism destination. *International Journal of Tourism Research*, 12(2), 116–133. <https://doi.org/10.1002/jtr.738>
- Smith, R. (2023, September 16). Using branded documentaries in modern marketing strategies. Roo Smith. <https://roosmith.com/blog/using-branded-documentaries-in-modern-marketing-strategies>
- Sofyan, A., Yuniati, Y., & Ratnasari, A. (2023). Copywriting marketing in reaching the target market. *KnE Social Sciences*. <https://doi.org/10.18502/kss.v8i18.14246>
- Sugarman, J. (2007). *The Adweek copywriting handbook: The ultimate guide to writing powerful advertising and marketing copy from one of America's top copywriters*. John Wiley & Sons. <https://www.wiley.com/en-us/The+Adweek+Copywriting+Handbook%3A+The+Ultimate+Guide+to+Writing+Powerful+Advertisin+g+and+Marketing+Copy+from+One+of+America's+Top+Copywriters-p-9780470051245>
- Wibowo, F. (2007). *Teknik produksi program televisi*. Pinus Book Publisher. <https://perpustakaan.binadarma.ac.id/opac/detail-opac?id=17303>
- Wu, G., & Ding, X. (2023). Which type of tourism short video content inspires potential tourists to travel? *Frontiers in Psychology*, 14, 1086516. <https://doi.org/10.3389/fpsyg.2023.1086516>
- Zabukovec Baruca, P., & Čivre, Ž. (2022). Unique destination attributes as a basis of tourism experience. *Academica Turistica*, 15(3), 349–362. <https://doi.org/10.26493/2335-4194.15.349-362>
- Zheng, Z., Lu, C., & Zhao, W. (2023). Study on the influence of Douyin short video marketing on camping travel intention. *SHS Web of Conferences*, 179, 03022. <https://doi.org/10.1051/shsconf/202317903022>
- Zulkifly, H. Z., & Firdaus, N. (2014). Persuasion and the online consumers: Investigating copywriting strategies in native advertisements. *International Journal of Social Science and Humanity*, 4(6), 430–434. <https://doi.org/10.7763/IJSSH.2014.V4.393>

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright ©2025 Dhini Aulia, Ahmad Bukhori Muslim, Budi Hermawan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Involving students in a collaborative and reflective reading material development: A participatory action research study

Elok Putri Nimasari,¹ Bambang Harmanto,¹ Rinjani Bonavidi,² Adi Fajaryanto Cobantoro,¹ Elvina Dwi Cahyani,¹ Velrinda Fahza'a Risty¹

¹Universitas Muhammadiyah Ponorogo, Indonesia, ²Keithcot Farm Children Centre Department for Education South Australia, Australia

Grounded in Halliday's Systemic Functional Linguistics and Freire's Critical Pedagogy, this participatory action research aims to develop a transformative digital reading assessment model that embeds cultural values and promotes critical literacy. Conducted over four months at an Islamic-Muhammadiyah higher education institution in Indonesia, this study involved 30 student participants and several community figures. Data were collected through classroom observations, focus group discussions, and interviews, and analyzed using critical discourse analysis informed by SFL's meta functions and Freirean principles. Findings reveal three key themes: (1) students interpreted texts as ideological constructs by identifying field, tenor, and mode; (2) students co-constructed culturally situated reading materials through dialogic engagement with community members; and (3), students critically engaged with multimodal digital texts by analyzing how visual and linguistic elements shaped meaning. This study offers a practically grounded model for reading assessment, showing how digital platforms can serve as ethical spaces for language learning that is both critical and culturally embedded.

Keywords: critical pedagogy, digital reading assessment, cultural values, systemic functional linguistics

OPEN ACCESS

ISSN 2503 3492 (online)

*Correspondence:

Elok Putri Nimasari
nimasari@umpo.ac.id

Received: 16th June 2025

Accepted: 14th August 2025

Published: 8th October 2025

Citation:

Nimasari, E.P., Harmanto, B., Bonavidi, R., Cobantoro, A.F., Cahyani, E.D., Risty, V.F. (2025). *Involving students in a collaborative and reflective reading material development: A participatory action research study*. *JEES (Journal of English Educators Society)*, 10(2). <https://doi.org/10.21070/jees.v10i2.1948>

INTRODUCTION

Reading assessment in English for Specific Purposes (ESP) has predominantly focused on cognitive processing models and standardized testing approaches, often overlooking the socio-cultural and ideological dimensions of language embedded in reading texts (Basturkmen, 2022; Hyland, 2019). Michael Halliday's Systemic Functional Linguistics (SFL), however, provides a powerful framework for examining how texts function as social semiotic resources that both shape and are shaped by discourse communities (Halliday & Matthiessen, 2014). SFL posits that texts operate within three interrelated dimensions of meaning: field (the subject matter and its social significance), tenor (the relationships and power dynamics between language users), and mode (the channel of communication and its degree of formality). These dimensions are crucial for contextualizing reading assessment, particularly when embedding specific ideological and cultural values, such as Muhammadiyah values, into digital learning environments in higher education.

Despite the growing body of literature in ESP and digital learning, studies integrating Halliday's SFL into reading assessment remain limited (Dreyfus et al., 2021; Ryshina-Pankova & McKnight, 2022). More specifically, little is known about how reading comprehension materials in digital learning environments can be structured to embed Islamic cultural values while ensuring that texts remain functionally meaningful and contextually relevant.

Traditional reading assessments tend to prioritize linguistic accuracy and comprehension strategies, yet they often overlook how texts construct social identities and ideological positions (Martin & Rose, 2022). Addressing this gap, an SFL-informed approach to reading assessment in higher education can strengthen instructional materials by aligning textual construction with the cultural values and ideological orientations central to Islamic, particularly Muhammadiyah, educational philosophy. Muhammadiyah, one of the largest Islamic organizations in Indonesia, was founded in 1912, with a mission that emphasizes education, social welfare, and the promotion of progressive Islamic thought. Its values emphasize modernity, rationality, and ethics rooted in Islamic teachings, all of which play a significant role in shaping the educational philosophy of Muhammadiyah institutions (Wijayati & Khafidhoh, 2023).

Recent studies have explored the application of SFL in language education, particularly in writing assessment (Derewianka & Jones, 2020; Moradian et al., 2021) and classroom discourse analysis (Fang & Schleppegrell, 2020; Jackson, 2021). In contrast, research on ESP reading has been slower to adopt SFL as a core analytical tool. Studies by Maton and Doran (2021), Mochizuki and Starfield (2021), and Morton (2023) highlight the role of semantic waves in structuring academic texts, while Hyland (2021) and Leezenberg (2002) discuss disciplinary variation in language choices. However, these works focus primarily on academic writing rather than reading assessment. Furthermore, digital learning environments have increasingly incorporated corpus linguistics and automated assessment tools (Alaqlobi et al., 2024; Godwin-Jones, 2020; Van Der Vegt et al., 2023), yet these approaches often overlook the social and functional dimensions of language use (Ando, 2021; Delahunty et al., 2014). Although digital platforms now offer interactive and multimodal reading materials (Al-Mahrooqi & Denman, 2022; Nimasari et al., 2023; Silseth & Gilje, 2019; Weinmann et al., 2024), there remains limited research on how such materials can be critically designed to embed specific cultural values and foster ideological literacy through language choices aligned with SFL principles.

The core issue in current reading assessment practices for ESP lies in their neglect of the social and ideological functions of texts. Traditional assessments tend to emphasize lexical and grammatical accuracy while overlooking how meaning is constructed in relation to the social purposes of communication (Halliday & Matthiessen, 2014). Although existing approaches, such as genre-based pedagogies (Hyland, 2021) and discourse-based assessments (Danis, 2022), offer valuable insights, they seldom address the ideological dimensions of reading texts, particularly within religious and cultural education contexts like Muhammadiyah higher education institutions. This gap is especially significant in the context of Indonesian higher education, where cultural values shape students' perspectives, ideological literacy fosters critical awareness of belief systems, and critical literacy empowers learners to question, analyze, and transform texts. Addressing this intersection of language, ideology, and culture is therefore essential for developing more socially responsive and contextually grounded reading assessment models.

A growing body of research suggests that embedding cultural and ideological values in reading materials can enhance learner engagement and foster critical literacy (Li et al., 2023; Luke & Freebody, 2021; Li et al., 2024; Zhang & Wang, 2024). However, few studies have examined how these values can be explicitly incorporated into digital reading texts through the use of SFL as both an analytical and pedagogical framework. Integrating Muhammadiyah cultural values into digital reading assessments requires a systematic approach to field (content relevance), tenor (ideological positioning), and mode (digital affordances and multimodal communication). Moreover, integrating critical pedagogy, particularly as conceptualized by Freire (1970), can strengthen this framework by ensuring that reading assessments move beyond measuring language proficiency to fostering critical engagement with texts and ideological structures they convey (Corbett & Guilherme, 2021). Critical pedagogy promotes critical consciousness, enabling students to reflect on and challenge dominant narratives in reading materials.

This study seeks to address this gap by developing and accessing reading materials that embed Muhammadiyah values within an SFL framework, thereby offering a model for contextually enriched digital reading assessments in higher education. Specifically, it examines how students engage in a collaborative and reflective reading material development using Halliday's field, tenor, and mode framework while integrating Freire's critical pedagogy with its emphasis on student agency and critical consciousness.

The novelty of this research lies in its interdisciplinary approach, combining SFL (Halliday & Matthiessen, 2014) and critical pedagogy (Freire, 1970; Corbett & Guilherme, 2021; Freire & Macedo, 2014) to advance ESP reading assessment. By incorporating Halliday's linguistic theory (Halliday & Matthiessen, 2014; Pakir, 2019) into digital learning materials and fostering critical engagement with texts, this study offers a perspective on reading assessment that is both linguistically rigorous and socially transformative. This research contributes empirically by providing evidence-based insights into how SFL and critical pedagogy can be operationalized in ESP reading assessment. Practically, it offers educators in Muhammadiyah institutions a structured framework for embedding ideological values in digital reading materials, thereby promoting a more holistic and contextually responsive approach to language education.

Gaps from Previous Studies

In recent years, numerous studies have applied SFL in language education, particularly in discourse analysis, literacy development, and pedagogical interventions. However, substantial gaps remain in the integration of SFL into reading assessment, especially in embedding ideological and cultural values, such as Muhammadiyah principles, within digital learning frameworks. This review critically examines relevant studies from Scopus-indexed journals, summarizing their aims, methodologies, and key findings to identify research gaps that establish the novelty of the present study.

Ryshina-Pankova et al. (2021) provided a comprehensive overview of SFL as a framework for analyzing multilingual and multimodal discourse. Their study focused on the ideational, interpersonal, and textual meta functions in

multilingual literacy education, highlighting how SFL enables learners to navigate academic texts effectively. Using qualitative discourse analysis, the study examined instructional strategies aligned with SFL principles. While the findings affirmed SFL's value as a powerful analytical tool in multilingual classrooms, the research did not extend its application to reading assessments or the embedding of ideological content within texts. Thus, although it provided important pedagogical insights, it lacked empirical data on assessment methodologies that integrate ideological values, a gap this study aims to address.

[Danis \(2022\)](#) and [Thomas \(2022\)](#) examined writing assessment through the lens of SFL, emphasizing that meaning making should be a central to language testing. Using corpus linguistics to analyze student writing samples, their studies demonstrated how textual choices reflect social positioning and disciplinary knowledge. However, their scope was confined to writing rather than reading assessments, leaving unexplored how SFL can be applied to evaluate comprehension skills in digital learning environments. Moreover, neither study address the influence of ideological and cultural values on reading comprehension, underscoring the need for further research in this area.

[Suwannasom and Pattanajak \(2023\)](#) investigated the role of digital learning ecosystems in higher education, focusing on classroom teaching methodologies in Thailand. Their mixed-method study identified key components of digital learning environments, including interactive learning materials, instructional support, and collaborative platforms. While offering a detailed analysis of digital pedagogy, the study did not consider the linguistic structuring of reading materials through an SFL framework. Additionally, it overlooked how ideological values could be systematically embedded in digital texts, leaving an unexplored intersection between digital learning, linguistic theory, and ideological content.

[Morton \(2023\)](#) explored instructional design in higher education through an SFL lens, emphasizing the alignment of learning outcomes with genre-based pedagogy. Using action research methods, this study developed instructional frameworks grounded in genre theory and SFL principles. While the findings supported the effectiveness of SFL in structuring instructional design, the research did not extend to digital reading assessments or the integration of cultural and ideological values. This limitation indicates a gap in research that explicitly examines how SFL-informed reading materials can be developed to reinforce cultural identity and ideological perspectives within digital learning platforms.

A broader review by [Alaqlobi et al. \(2024\)](#) and [Al-Mahrooqi and Denman \(2022\)](#) assessed the strategic integration of artificial intelligence in higher education, emphasizing how AI-driven digital tools enhance learning experiences. While both studies acknowledged the potential of technology for personalized learning and assessment, they lacked a linguistic focus and did not incorporate SFL in their analysis of digital texts. Additionally, they did not explore how digital reading assessments could be designed to embed specific cultural or ideological narratives, another gap this study seeks to address.

Despite the valuable contributions of these studies to language education, several critical gaps remain unaddressed. First, none explicitly examined reading assessments from an SFL perspective, particularly in relation to embedding ideological values in digital learning materials. While some research explored genre-based pedagogies and digital learning ecosystems, they did not consider how Muhammadiyah cultural values could be systematically integrated into reading texts through the dimensions of field, tenor, and mode. Additionally, prior studies have predominantly focused on writing or general literacy education, leaving reading comprehension assessment as an underexplored area in SFL-based research.

Another critical gap is the lack of empirical studies at the intersection of digital learning, ideological content, and functional grammar ([Danis, 2022](#); [Godwin-Jones, 2020](#)). Although research has examined digital learning platforms or linguistic theory separately, there remains a significant void in studies that integrate these aspects to enhance ESP reading assessment. Moreover, existing studies have not proposed practical, empirically tested frameworks for developing and accessing digital reading materials that incorporate ideological and cultural dimensions.

To answer these gaps, this study proposes a novel approach to reading assessment that integrates SFL principles with Muhammadiyah cultural values in digital learning environments. By analyzing and designing reading materials through Halliday's dimensions of field, tenor, and mode, this study aims to develop an assessment framework that not only evaluates comprehension skills but also fosters ideological engagement. This interdisciplinary approach bridges linguistic theory, digital pedagogy, and cultural education, offering a comprehensive model for ESP reading assessment in higher education. The study will contribute to the literature by providing empirical evidence on how digital reading assessments can be purposefully designed to embed cultural values, promote critical awareness, and enhance learning outcomes.

Theoretical framework

Halliday's systemic functional linguistics (SFL) offers a robust theoretical foundation for developing reading assessment frameworks that move beyond traditional emphases linguistic accuracy and cognitive processing. As a social semiotic system, SFL views language as a resource for meaning making within specific social and ideological contexts ([Halliday & Matthiessen, 2014](#); [Pakir, 2019](#)). This study applies SFL's three meta-functions, ideational, interpersonal, and textual, to the development of reading texts and assessments, ensuring that materials are both comprehensible and contextually meaningful.

To enhance the socio-critical dimension, SFL is integrated with [Freire \(1970\)](#) critical pedagogy ([Corbett & Guilherme, 2021](#)), which positions education as a means of social transformation and promotes critical consciousness through questioning dominant discourses. The ideational meta-function supports the analysis of how texts represent social reality, while critical pedagogy ensures that students interrogate the ideological structures underpinning these representations. The interpersonal meta-function aligns with

Freire’s dialogic learning, fostering co-construction of knowledge between educators and students. Using participatory action research (PAR), this study incorporates student perspectives into digital reading assessment design to promote inclusivity and agency. Finally, the textual meta-function ensures coherence and cohesion in digital reading texts, while critical pedagogy frames literacy as a tool for empowerment. Digital platforms will be designed to encourage active, multimodal engagement, enhancing both linguistic proficiency and critical literacy.

By integrating SFL with critical pedagogy, this study proposes a transformative reading assessment framework that not only evaluates linguistic proficiency but also fosters ideological awareness and critical thinking in digital learning environments. SFL offers a systematic approach to analyzing language in context, while critical pedagogy offers a means to interrogate the power relations and ideological positions embedded within reading texts (Freire, 1970; Corbett & Guilherme, 2021; Freire & Macedo, 2014). This integration ensures that digital learning environments do not merely provide linguistic input but also engage students in an active process of meaning-making and critical reflection.

In the participatory action research (PAR) design, learners are positioned as active agents in the learning process rather than passive recipients of knowledge (Leavy, 2023). Freire’s concept of dialogic learning is operationalized through student involvement in the assessment process, allowing them to reflect on their understanding of texts and their ideological implications (Corbett & Guilherme, 2021; Freire & Macedo, 2014). This approach moves reading assessments beyond static comprehension questions, fostering deeper engagement with language as a tool for social and ideological critique.

Furthermore, digital multimodal texts expand the possibilities for applying SFL and critical pedagogy in reading assessment. Texts in digital learning environments extend beyond linear print to include images, hyperlinks, videos, and interactive elements. These multimodal resources enable students to engage with multiple layers of meaning (Nadia et al., 2021), reinforcing SFL’s textual meta-function while simultaneously fostering critical engagement with digital literacy practices (Ryshina-Pankova et al., 2021). Through this framework, reading assessments can challenge students to recognize how meaning is shaped not only by linguistic structures but also by multimodal affordances. This study aims to bridge the gap between functional linguistics and socio-critical approaches, offering a holistic and participatory model for developing reading assessments that reflect both language and ideology. By embedding Islamic-

Muhammadiyah cultural values into digital reading texts through the combined application of SFL and critical pedagogy, this research contributes to the growing discourse on integrating ideological and cultural perspectives in higher education reading assessment.

METHODS

Research Context and Participants

Grounded in social-constructivism research paradigm (Leavy, 2023; Pass, 2007), this four-month participatory action research project was conducted at an Islamic-Muhammadiyah higher education institution in East Java, Indonesia. The study explored digital learning environments to develop reading materials that embed Islamic values. Of the 54 eligible students, 30 were invited to participate voluntarily, selected based on their interest in creating culturally embedded reading materials and their willingness to engage in reflective, dialogic classroom activities. Before selection, all students received a detailed briefing on the project’s objectives, expectations, and time commitments. Those who provided informed consent joined the PAR process as a co-researcher. No participants were assigned; involvement was entirely voluntary to ensure authentic engagement and alignment with the emancipatory aims of Freirean pedagogy.

The researchers assumed on multiple roles in this study, serving both as lecturers to the student participants and as observers of the material development process. Alongside the students, five community figures, two from Muhammadiyah and three from Aisiyiyah, participated as research collaborators. These individuals were purposively selected for their extensive historical and ideological knowledge of Muhammadiyah and Aisiyiyah values, as well as their active engagement in local cultural education. Their contributions included co-developing content, providing historical artifacts, and participating in iterative discussions with both students and researchers. Data were collected through observations and discussions with all participants. Initial meetings were conducted to obtain informed consent and clarify participant roles in the project. The researchers met with the Muhammadiyah and Aisiyiyah collaborators on five occasions to collect and analyze historical data. Meanwhile, the student participants attended six reading comprehension classes, where they engaged in critical reading activities based on the texts under study.

TABLE 1 | Summary of PAR stages, timeline, and participant roles

PAR’s stages	Timeline	Researcher Roles	Student Roles
Stage 1: Orientation and Planning	Week 1–2	Provided demonstrations; introduced the SFL and Freirean frameworks; obtained informed consent.	Participated in orientation; selected topics for reading; discussed the relevance of texts.
Stage 2: material collection	Week 3–5	Facilitated meetings with Muhammadiyah/Aisiyiyah figures; observed interactions.	Conducted interviews with community figures; gathered historical materials.

Stage 3: collaborative analysis	Week 6–8	Guided group discussions; observed student interactions and interpretations.	Analyzed texts using the SFL framework; interpreted ideological content; co-constructed texts.
Stage 4: Reflection and Revision	Week 9–12	Conducted focus groups and interviews; documented student reflections.	Reflected on engagement; revised materials; developed digital reading tasks.

As the project progressed, student participants were given the autonomy to select their own reading topics through structured group discussions. To ensure authenticity and contextual relevance, they engaged directly with Muhammadiyah and Aisyiyah figures, gathering primary sources and historical artifacts. The collected materials were then brought back to the classroom for peer discussions, where students collaboratively analyzed the texts. This extensive reading activity enabled them to engage more deeply with the content and construct meaning that was both personally and academically significant.

Throughout the project, students were encouraged to self-monitor their reading progress on a weekly basis. This reflective practice allowed them to track their own development, assess comprehension and analytical skills, and adjust reading strategies accordingly. Consistent self-monitoring fostered greater autonomy, heightening students' awareness of their learning processes and the progression of their reading abilities over time.

Data Collection and Analysis

This PAR was conducted in an ESP class at an Islamic-Muhammadiyah higher education institution in Indonesia, engaging 30 eligible students as co-researchers. Data were gathered from multiple sources to enable triangulation and gain a deeper understanding of participants' involvement in the collaborative and reflective development of reading material. These sources included participant observations, focus group discussions, and semi-structured interviews.

Research Procedures

At the outset of the PAR project, the first author, who also served as lecturer and research collaborator, provided demonstrations to guide students in selecting appropriate reading texts (see [Table 1](#)). Students were introduced to text analysis through the lens of Systemic Functional Linguistics (SFL) and critical pedagogy, enabling them to explore how language functions within a social context. Particular emphasis was placed on identifying socio-cultural values and underlying messages embedded in texts. This stage established a foundational framework for students to critically engage with reading materials and recognize how texts convey meaning beyond their surface content.

Throughout the intervention, the first author, acting as a teacher-researcher, closely observed classroom interactions, group work, and student engagement. These participant observations captured the dynamics of collaborative text-making as well as moments of reflective dialogue and critical questioning. Focus group discussions were conducted with two different sets of stakeholders. The first involved ongoing discussions between the researcher and students, focusing on how reading materials were jointly developed, reflected upon, and revised. These sessions served as dialogic spaces in

which students could articulate understandings, negotiate meaning, and raise critical inquiries into texts and contexts. The second set of focus groups included key community figures affiliated with Muhammadiyah and Aisyiyah, contributing broader cultural and ideological contextualization to the materials under development.

Additionally, semi-structured interviews were conducted with Muhammadiyah and Aisyiyah figures to explore how Islamic pedagogical values could inform – or challenge – the textual and discursive features of the collaboratively produced materials. These conversations offered critical insights into how ideological stances influenced students' engagement and decision-making during the material development process. All interactions, including classroom observations, focus groups, and interviews – were audio-recorded and transcribed verbatim. The resulting transcripts constituted the primary corpus of qualitative data, which was analyzed through critical discourse analysis informed by Halliday's SFL and Paulo Freire's critical pedagogy (see [Table 2](#)).

From an SFL perspective, the transcribed texts were analyzed with attention to the ideational meta-function (what is represented), the interpersonal meta-function (how relationships are enacted), and the textual meta-functions (how texts are organized). This analytical lens enabled the researcher to examine how meanings were constructed and negotiated within the texts co-authored by students, and how linguistic choices either reinforced or resisted dominant discourses. Concurrently, the principles of Freirean critical pedagogy were applied to interrogate how students engaged with texts as active subjects of their own learning. Particular attention was directed to moments in which students challenged taken-for-granted assumptions, questioned power relations embedded within texts, and collectively re-authored materials to more accurately reflect their social realities.

Through iterative cycles of coding, categorization, and interpretation, emergent themes were identified. These themes captured the progression of students from passive recipients of textual knowledge to critical co-creators of culturally resonant and pedagogically meaningful texts (see [Table 2](#)). The integration of SFL and critical pedagogy in the analysis provided a dual lens: one structural and linguistic, the other ideological and transformative. Credibility was strengthened through methodological triangulation across observations, focus groups, and interviews, as well as through member checking with participants to validate emerging interpretations. To mitigate potential bias, the researcher maintained reflexive field notes and engaged in peer debriefing. Given the ideological sensitivity of the content, the researcher's positionality as both facilitator and co-participant was critically examined to ensure transparency and analytic rigor.

TABLE 2 | Thematic coding analysis from Hallidayan and Freirean theories

Theoretical framework	Main themes	Sub-themes	Illustrative analysis examples
Hallidayan Systemic Functional Linguistics (SFL)	Interpreting texts as ideological constructs through SFL	<ul style="list-style-type: none"> Ideational meta-function (field) Interpersonal meta-function (tenor) Textual meta-function (mode) 	<ul style="list-style-type: none"> Identifying value-laden language (e.g., “KH. Ahmad Dahlan established...”) Recognizing authoritative tone and relational distance Analyzing chronological sequencing to frame professionalism and vision
Freirean Critical Pedagogy (dialogue, conscientization) and Hallidayan SFL (structuring text via field, tenor, mode)	Dialogic co-construction of culturally situated materials	<ul style="list-style-type: none"> Collaboration with community figures Reflective selection of topics Ideological negotiation in text development 	<ul style="list-style-type: none"> Inclusion of gender perspectives (e.g., Text 3: Aisyiyah) Emphasis on social justice (e.g., Al-Ma’un and economic empowerment) Students designing questions targeting field, tenor, and mode
Hallidayan SFL (textual/meta-functional analysis of multimodality) and Freirean Critical Pedagogy (critical questioning, ideological critique)	Critical engagement with digital and multimodal texts	<ul style="list-style-type: none"> Multimodal text analysis Ideological interpretation of visual/linguistic synergy Student-created digital reading tasks 	<ul style="list-style-type: none"> Analyzing visual-textual alignment (e.g., image placement & hyperlinks) Evaluating ideological implications in Qur’anic verse sequencing Designing comprehension tasks addressing digital mode and layout choices

RESULTS AND DISCUSSION

Based on selective data analysis from classroom observations, student focus group discussions, and interviews with Muhammadiyah and Aisyiyah figures, three central themes emerged, reflecting the theoretical integration of Halliday’s SFL (Halliday & Matthiessen, 2014; Pakir, 2019) and Freire’s critical pedagogy (Corbett & Guilherme, 2021; Freire, 1970; Freire & Macedo, 2014). These themes demonstrate how students, positioned as both co-authors and critical readers, actively engaged in developing ideologically enriched reading materials within digital learning environments. The three themes are: (1) interpreting texts as ideological constructs through SFL; (2) dialogic co-construction of culturally situated materials; and (3) critical engagement with digital and multimodal texts.

Interpreting Texts as Ideological Constructs through SFL

Students developed critical awareness of how language functions to represent, position, and structure social reality. Through the lens of SFL’s ideational, interpersonal, and textual meta-functions, they analyzed reading texts to uncover how meaning is shaped by field (what is being discussed), tenor (who is involved and in what relationship), and mode (how the message is delivered). This theme highlights how the reading assessment process extended beyond the mere decoding of information to a deeper interrogation of whose values and ideologies were embedded within the text. Students’ ability to both recognize, and re-author texts reflects SFL’s effectiveness as a meaning-making

model and its potential for fostering ideological literacy (Gebhard & Accurso, 2022; Martin & Rose, 2022; Morton, 2023).

Classroom engagement with Text 1: The Founding of Muhammadiyah offers a clear example of how students deconstructed ideational meaning. They discussed how the text frames KH. Ahmad Dahlan’s establishment of Muhammadiyah as both a religious and revolutionary act. Several participants pointed out the recurrent use of material and relational processes, such as “KH. Ahmad Dahlan established the school...” and “Muhammadiyah was officially recognized...”, which serve to convey purposeful agency and institutional legitimacy rooted in Islamic reform. In a focus group session, the students stated:

“When the text says, ‘He relied on his own wealth to realize his vision of a modern Islamic educational institution,’ I think it tries to show KH. Ahmad Dahlan not only as a religious figure, but also as someone modern, rational, and generous. It’s not just telling history but promoting Muhammadiyah’s values.” (Student focus group transcript, FG1)

“I realized that the phrase ‘Muhammadiyah was officially recognized’ isn’t neutral, it makes it sound formal and state-approved, which adds to its legitimacy.” (Student FGD, FG2)

“This sentence, ‘with the hope that its members would follow the example,’ sounds like it’s talking down to the reader a bit. It makes us feel like followers, not participants.” (Student discussion, Session 3)

“Why not say, ‘They worked together to build schools’? It would sound more collective than just saying ‘KH. Ahmad Dahlan established...’” (Classroom reflection, Session 2)

These students’ excerpts reflect their understanding of the field dimension of SFL (what is being represented) and how it aligns with Muhammadiyah’s ideology of Islamic modernism and social reform. In another observation, students noticed how the tenor of the text, marked by respectful references to KH. Ahmad Dahlan and the use of a collective narrative voice (“with the hope that its members would follow the example...”), served to reinforce the organization’s authority and moral appeal. Additionally, the textual meta-function emerged in students’ analysis of the text’s chronological sequencing, from its founding to its expansion, which they interpreted as not merely informative but also persuasive in positioning Muhammadiyah’s growth as a logical and inevitable progression. One student remarked:

“The structure makes Muhammadiyah look organized and forward thinking. It helps us see that it’s not only spiritual but also professional.” (Student observation note, Session 3)

Interview with Muhammadiyah and Aisiyiah figures revealed a strong consensus on the importance of including The Founding of Muhammadiyah as a foundational reading text. One Muhammadiyah elder explained:

“Our young generation must understand that Muhammadiyah was born from struggle and vision. It’s not just a name. If they know the history, they will realize why values like education, rationality, and social welfare matter. This is essential for shaping them as agents of change.” (Interview, Muhammadiyah representative, IM1)

Another Aisiyiah figure emphasized that embedding this text helps bridge identity and purpose:

“Reading about Ahmad Dahlan’s efforts helps students connect personally with our mission. They need to see that they are continuing a legacy, not just reading for a test.” (Interview, Aisiyiah leader, IA2)

These findings validate the decision to foreground The Founding of Muhammadiyah within the assessment model and illustrate how history operates ideologically through SFL’s field (subject relevance), tenor (relationship-building), and mode (chronological, formal presentation). The text’s cultural relevance not only increased student engagement but also reinforced their role as moral and intellectual inheritors of Muhammadiyah’s vision. (see [Figure 1](#))

While prior research by [Ariyanti \(2023\)](#) and [Nimasari, et al. \(2024\)](#) has examined the application of SFL in writing assessment and instructional design, limited attention has been given to its use in reading assessment or the integration of ideological values into digital learning materials. The distinctiveness of this study lies in its shift from grammar-focused comprehension toward a meaning-making approach that synthesizes SFL with [Freire’s \(1970\)](#) principles of critical pedagogy ([Corbett & Guilherme, 2021](#)). In contrast to automated or corpus-based reading tasks ([Godwin-Jones, 2020](#)), the student-led analyses in this project were both context-aware and culturally embedded.

Compared to earlier studies that treat functional linguistics and ideological critique as separate domains, this

research demonstrates that SFL can serve as a powerful analytical framework for revealing how reading texts function as vehicles for ideological formation. By explicitly embedding Muhammadiyah cultural values into the dimensions of field, tenor, and mode, and by equipping students to analyze them, the study advances critical literacy while situating it within a local Islamic framework. Moreover, the integration of student voice and agency aligns with Freire’s vision of learners as active participants in the construction of knowledge. The transformative shift was evident when students moved beyond asking, “What is this text about?” to posing questions such as “Whose perspective is this?” and “What values are being promoted?”

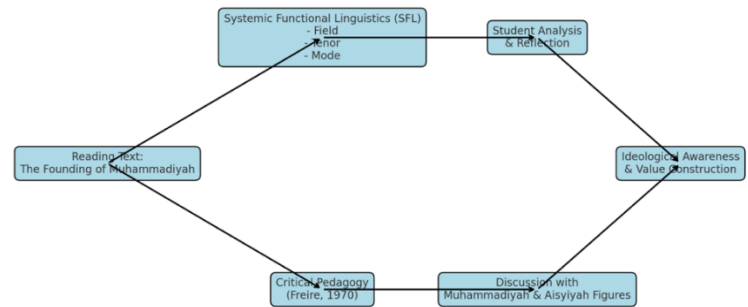


FIGURE 1 | Visual Model: Interpreting Texts as Ideological Constructs through SFL

Dialogic Co-Construction of Culturally Situated Materials

Aligned with Freire’s emphasis on dialogic learning, this theme demonstrates how students and community figures collaboratively developed reading materials that embodied Islamic and Muhammadiyah cultural values. Rather than relying on pre-designed resources, students actively gathered, analyzed, and curated materials drawn from authentic artifacts and local narratives. These iterative cycles of negotiation allowed for the integration of SFL-informed linguistic structuring with ideological reflexivity, transforming the classroom into a participatory site of cultural production ([Luke & Freebody, 2021](#); [Nimasari et al, 2024](#); [Li et al., 2024](#)). In this process, learners were positioned as producers of knowledge rather than passive recipients, reflecting Freire’s vision of education as a liberatory practice.

During collaborative discussions and material development stages, students engaged directly with Muhammadiyah and Aisiyiah figures to curate texts of historical, ideological, and cultural significance. One of the clearest examples of this co-construction process emerged in the selection and development of Text 3: Aisiyiah. Initially, many students overlooked this text in favor of more prominent figures and events. However, following a dialogue with an Aisiyiah leader, the importance of its inclusion became evident. The leader explained:

“Aisiyiah has been shaping Islamic education for women long before many of us were born. If students only read about male figures, how will they understand the gender values Muhammadiyah stands for?” (Interview, IA3)

In response, students collaboratively drafted comprehension questions, discussion prompts, and sentence completion tasks centered on the Aisiyiah narrative,

reflecting their growing understanding of gender representation and ideological discourse. This process aligns with Freire’s concept of “conscientization”, the development of critical awareness through reflection and dialogue (Freire, 1970). Another powerful example emerges from Text 4: Muhammadiyah Economic Movement, which was developed at the request of a Muhammadiyah figure who urged:

“Economic empowerment is part of our da’wah. Students need to read about Al-Ma’un not just as charity, but as a call to change society structurally.” (Focus group discussion, Muhammadiyah leader, FG-M1)

This insight prompted students to revise the original draft to emphasize ideational elements related to collective economic action and social justice. They incorporated questions addressing the field (economic responsibility), tenor (Islamic moral authority), and mode (informational sequencing), thereby reinforcing all three meta-functions of SFL. An excerpt from a classroom transcript illustrates this co-construction in practice:

Student A : *“Why do we think Aisyiyah is less represented in textbooks?”*

Student B : *“Maybe because people think it’s just a women’s group, but actually it started formal schooling for girls, right?”*

Student C : *“Yeah, and Nyai Ahmad Dahlan was also a founder, why don’t we highlight that?”*

(Classroom transcript, Session 4)

“We only see men in our textbook stories. But what about Nyai Ahmad Dahlan? Why didn’t we include her in the reading? She’s as important.” (Student comment, Session 4)

This transcription from classroom discussion session 4 demonstrates how students critically evaluated dominant narratives and restructured content to challenge gender marginalization in institutional history, aligning with SFL’s interpersonal function and Freirean dialogic pedagogy. Previous research, such as Zhang and Wang (2024) and Suwannasom and Pattanajak (2023), has emphasized the utility of SFL for instructional design and digital materials. However, these studies rarely position students as co-developers of materials, nor do they involve community figures in validating or enriching ideological content. The integration of community-based dialogue into ESP materials therefore represents an innovation that has not yet been sufficiently emphasized in the current literature.

The dialogic material development process identified in these findings contributes to the field in several keyways. It operationalizes critical pedagogy and SFL not only in text analysis, but also in authorship and curriculum design. It fosters learner autonomy, critical thinking, and engagement with authentic sociocultural knowledge. It bridges formal academic literacy with community knowledge, thereby enhancing content relevance and ideological depth. Unlike previous ESP assessments that focus on abstract or generic topics, the texts produced in this study, such as Aisyiyah, Buya Hamka, and The Muhammadiyah Economic Movement, are imbued with values, histories, and ideologies that reflect students’ identities and social contexts.

Furthermore, these findings demonstrate that dialogic co-construction allowed students to act as cultural agents and ethical authors. They did not merely complete a prescribed learning task; rather, they participated in shaping educational discourse, transforming the classroom into a site for social change, fully aligned with Freire’s (1970) pedagogical vision and Halliday’s conception of language as a social semiotic (Jackson, 2021).

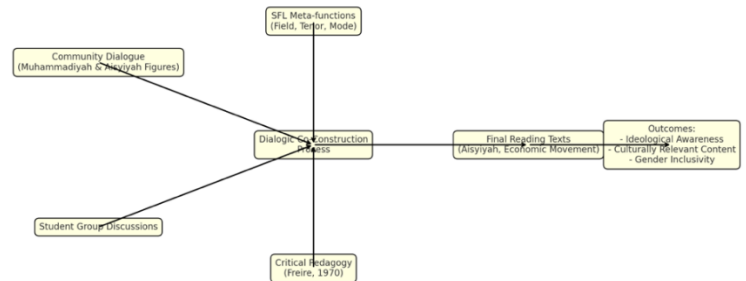


FIGURE 2 | Visual Model: Dialogic Co-Construction of Culturally Situated Materials

Critical Engagement with Digital and Multimodal Texts

Students’ engagement with digital reading materials revealed how multimodal resources, such as images, hyperlinks, timelines, embedded media, interact with linguistic choices to shape meaning. Using the SFL framework, students learned to analyze how textual and visual modes complement one another in construing field, tenor, and mode. Freirean critical pedagogy further supported this analytical process by prompting students to interrogate not only what is represented, but also what is omitted, distorted, or silenced. This theme illustrates the potential of digital learning environments to support critical literacy practices that are both language-oriented and socially grounded (Fang & Schleppegrell, 2020; Al-Mahrooqi & Denman, 2022).

Digital multimodal texts such as Text 2: Buya Hamka and Text 5: Muhammadiyah’s Stance on Interfaith Relations offered students opportunities to interpret meaning across linguistic and non-linguistic modes. They examined how embedded hyperlinks, thematic visuals, and text layout influenced their understanding of narrative tone and ideological positioning. For example, one group examined the following sentence in Text 2.

“Hamka once said, as scholars, we have sold ourselves to God, and we cannot be sold to anyone else.”

During a classroom discussion, a student observed:

“This quote is powerful, but when we click the link and see how it’s framed next to the fatwa controversy, it adds a layer, like he’s being remembered more for his political stance than his religious values.” (Student discussion transcript, Session 5)

This example illustrates student’s growing awareness of interpersonal meta-function in digital settings, particularly how layout and hyperlink placement can influence a text’s tone and focus points. In their analysis of Text 5, students also reflected on the dual use of Qur’anic citations:

“The text places ‘Lakum dīnukum wa-liya dīnī’ before ‘fa-istabiqū al-khayrāt’, I think it wants us to think of separation before cooperation. Even the order has ideology.” (Focus

group discussion, FG4)

Such insights demonstrate students' capacity to critically analyze textual meta-functions, recognizing that meaning is shaped not only by what is written, but also by how it is sequenced and presented in digital form, processes that construct particular ideological readings. Further evidence of this critical engagement emerged when students designed comprehension questions for Text 4: Muhammadiyah Economic Movement. They debated whether the visuals of industry and finance used in the digital layout served to enhance or, conversely, to distort the intended message of al-Ma'un:

"The image of Buya Hamka next to that political quote makes him look like an activist, not a religious leader. That changes how we read it." (Student digital reading discussion, Session 5)

"It says 'economic empowerment,' but shows luxury buildings. That's not what Al-Ma'un is about. It's about helping the poor, right?" (Student design critique, Session 6)

"If the images only show big buildings, students might think the movement is just about wealth. But Al-Ma'un also means helping others." (Classroom reflection, Session 6)

This critique reflects an evolving awareness of ideational meta-functions within multimodal contexts, particularly in how visuals co-create or limit textual meaning. While existing literature acknowledges the potential of digital tools to deliver content interactively (Al-Mahrooqi & Denman, 2022; Suwannasom & Pattanajak, 2023), most studies emphasize technical functionality over ideological literacy. Few, if any, examine how multimodal elements, such as image placement, hyperlink pathways, and color schemes, influence the ideological reception of reading texts. In contrast, this present research demonstrates that when SFL is integrated with critical pedagogy, digital texts become more than platforms for language practice; they function as ideological artifacts that students can critically deconstruct.

The findings on critical engagement with digital and multimodal texts yield three practical contributions. First, they indicate a shift in digital reading from passive consumption to active critique, prompting students to ask, whose voices are amplified or silenced through these multimodal designs? Second, they equip students to analyze meaning across semiotic modes, including color, imagery, and hyperlink structures. Third, they reinforce the textual meta-function in digital discourse, an aspect rarely operationalized in ESP reading assessments. (see Figure 3)

Compared to prior models that treat form and content as separate entities, this study advances a framework in which form is content, an insight consistent with Halliday's conception of language as a social semiotic and Freire's view of literacy as a political act. Students' engagement with multimodal texts transformed the reading process into an act of critical inquiry. Through guided SFL analysis and dialogic classroom interactions, they come to recognize that texts are not neutral and that digital tools, far from being passive carriers of information, actively shape meaning. These findings confirm the transformative potential of integrating SFL and critical pedagogy in digital ESP contexts, offering a

replicable model for fostering critical literacy in culturally embedded education.

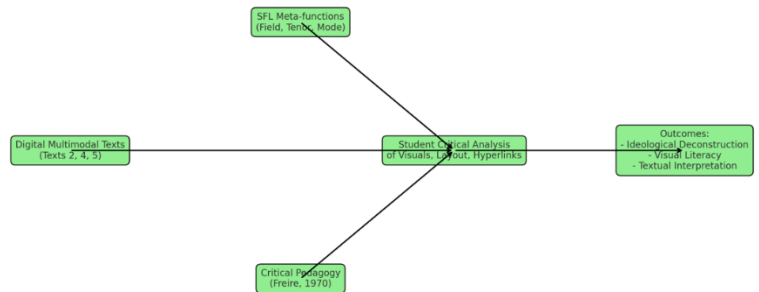


FIGURE 3 | Visual Model: Critical Engagement with Digital and Multimodal Texts

CONCLUSION

This study set out to develop and critically evaluate digital reading materials for ESP within an Islamic-Muhammadiyah higher education context by integrating Halliday's Systemic Functional Linguistics and Paulo Freire's critical pedagogy. Through a participatory action research approach, students were not merely learners but active co-authors, engaged in the co-construction of culturally relevant and ideologically embedded texts. The findings illustrate how the fusion of linguistic analysis and critical reflection can transform reading assessment into a dialogic, socially responsive practice.

One of the central findings is that students demonstrated a strong capacity to interpret texts as ideological constructs using the lens of SFL's three meta-functions, field, tenor, and mode. In analyzing texts such as "The Founding of Muhammadiyah" and "Buya Hamka," they identified how language is used to frame narratives, highlight moral perspectives, and construct social identities. Their interpretation moved beyond linguistic comprehension into the realm of ideological critique, underscoring the potential of SFL as a tool for fostering critical literacy.

Equally significant is the finding that the development of reading materials became a site for dialogic co-construction. Students collaborated with Muhammadiyah and Aisyiyah figures to select, adapt, and produce materials that reflected shared historical, cultural, and religious values. Through reflective classroom dialogue, they problematized issues such as gender representation, economic equity, and interfaith relations, revising texts to align more closely with Muhammadiyah's vision. This collaborative process not only deepened their engagement with the content but also positioned them as cultural and ideological authors, consistent with Freire's vision of transformative education.

The study also revealed the value of multimodal digital texts in facilitating critical engagement. Students were able to decode ideological meanings not only from written language but also images, hyperlinks, and text layouts. Their ability to analyze how the placement of Qur'anic verses or the sequencing of visuals influenced reader perception reflects an advanced level of critical digital literacy. This represents a significant shift from traditional print-based reading practices to a more dynamic, multimodal literacy model aligned with

the realities of contemporary digital learning environments.

Taken together, these findings offer two significant contributions. Theoretically, the research proposes a novel framework for ESP reading assessment that is culturally responsive, ideologically reflective, and digitally relevant. It challenges conventional paradigms that treat reading as a neutral cognitive act and instead positions it as a socially and politically situated practice. Practically, it offers innovations that can inform future curriculum development in Islamic and other culturally embedded education systems. Curriculum designers, policymakers, and ESP practitioners can adopt this model by integrating SFL and Critical Pedagogy, recognizing that language is used to construct meaning within diverse social contexts. Such an approach ensures that reading materials foster not only comprehension, but also critical awareness and cultural identity formation.

The study is not without limitations. Conducted within a single institutional context, the findings are context specific. Additionally, challenges in digital infrastructure, such as inconsistent internet access and varied levels of student digital literacy, may have affected the depth of engagement with the materials. Moreover, the study prioritized qualitative insights over standardized assessment metrics, leaving scope for future research to examine measurable impacts on learner outcomes.

Future studies might expand this work across a wider range of institutions, employ longitudinal designs to track the long-term effects of this framework on students' literacy and critical thinking, and explore the integration of SFL-based assessments into broader educational policies. Further investigation into the intersection of SFL, critical pedagogy, and technology across disciplines could also open new avenues for interdisciplinary education research.

In conclusion, this study highlights the transformative potential of integrating SFL with critical pedagogy in reading assessment practices. By embedding Muhammadiyah cultural values into digital learning texts and empowering students as co-creators of knowledge, it proposes a holistic, reflective, and context-sensitive approach to language education. This model not only supports linguistic development but also nurtures critical consciousness, affirming the role of language as a tool for personal growth, ideological engagement, and social transformation.

ACKNOWLEDGEMENTS

This research was funded by the Council for Higher Education, Research, and Development, Central Board of Muhammadiyah. The authors extend their sincere appreciation to the ESP students of Universitas Muhammadiyah Ponorogo, Pimpinan Daerah Muhammadiyah Ponorogo, and Ranting Aisyiyah Mangkujayan Ponorogo for their invaluable support and collaboration in the successful completion of this study.

REFERENCES

Alaqlobi, O., Alduais, A., Qasem, F., & Alasmari, M. (2024). Artificial intelligence in applied (linguistics): A

content analysis and future prospects. *Cogent Arts & Humanities*, 11(1), 2382422.

<https://doi.org/10.1080/23311983.2024.2382422>

Al-Mahrooqi, R., & Denman, C. (2022). Digital pedagogies in language education: A critical perspective. Routledge.

Ando, S. (2021). University teaching and learning in a time of social distancing: A sociocultural perspective. *Journal of Human Behavior in the Social Environment*, 31(1–4), 435–448.

<https://doi.org/10.1080/10911359.2020.1814928>

Ariyanti, A., Setiawan, S., & Munir, A. (2023). Students' enjoyment and anxiety in reminiscing about mind-mapping use in the English writing class. *Studies in English Language and Education*, 10(2), 789–804. <https://doi.org/10.24815/siele.v10i2.28215>

Basturkmen, H. (2022). Developments in English for specific purposes: A multi-perspective approach. Palgrave Macmillan.

Corbett, J., & Guilherme, M. (2021). Critical pedagogy and quality education (UNESCO SDG-4): The legacy of Paulo Freire for language and intercultural communication. *Language and Intercultural Communication*, 21(4), 447–454.

<https://doi.org/10.1080/14708477.2021.1962900>

Danis, N. (2022). A corpus analysis of disciplinary identity in evaluative journal articles: A Systemic Functional Linguistics approach. *English for Specific Purposes*, 68, 87–101. <https://doi.org/10.1016/j.esp.2022.07.003>

Delahunty, J., Verenikina, I., & Jones, P. (2014). Socio-emotional connections: Identity, belonging and learning in online interactions. A literature review. *Technology, Pedagogy and Education*, 23(2), 243–265. <https://doi.org/10.1080/1475939X.2013.813405>

Derewianka, B., & Jones, P. (2020). *Teaching language in context*. Oxford University Press.

Dreyfus, S., Humphrey, S., Mahboob, A., & Martin, J. R. (2021). *Genre pedagogy across the curriculum*. Equinox.

Fang, Z., & Schleppegrell, M. J. (2020). *Disciplinary literacies and systemic functional linguistics: Understanding texts in context*. Routledge.

Freire, P. (1970). *Pedagogy of the oppressed*. Bloomsbury.

Freire, P., & Macedo, D. P. (2014). *Pedagogy of the oppressed: 30th Anniversary Edition* (M. B. Ramos, Trans.; 30th anniversary ed.). Bloomsbury Publishing.

Gebhard, M., & Accurso, K. (2023). Pursuing a Multilingual Equity Agenda: SFL and Teacher Action Research in Hard Times. In *In Pursuit of a Multilingual Equity Agenda* (pp. 1–30). Routledge.

Godwin-Jones, R. (2020). Dealing with complexity in language learning: Automatic assessment, adaptive learning, and mobile-assisted language learning. *Language Learning & Technology*, 24(3), 1–13. <https://doi.org/10.125/44736>

Halliday, M. A. K., & Matthiessen, C. M. I. M. (2014). *Halliday's introduction to functional grammar* (4th ed.). Routledge.

Hyland, K. (2019). *Second language writing* (2nd ed.). Cambridge University Press.

- Hyland, K. (2021). *Genre and second language writing*. University of Michigan Press.
- Jackson, G. (2021). Axiological constellations in literary response writing: Critical SFL praxis in an ELA classroom. *Language and Education*, 35(5), 446–462. <https://doi.org/10.1080/09500782.2020.1856132>
- Leavy, P. (2023). *Research design: Quantitative, qualitative, mixed methods, arts-based, and community-based participatory research approaches* (2nd ed.). Guilford Press.
- Leezenberg, M. (2002). Power in communication: Implications for the semantics-pragmatics interface. *Journal of Pragmatics*, 34(7), 893–908. [https://doi.org/10.1016/S0378-2166\(01\)00067-4](https://doi.org/10.1016/S0378-2166(01)00067-4)
- Li, J., Wu, Y., & Gao, X. (2023). Language teachers' engagement with ideological representations as materials designers. *System*, 117, 103121. <https://doi.org/10.1016/j.system.2023.103121>
- Li, X., Islam, M. M., Islam, S., Gao, X., & Rong, Z. (2024). Examining professional ideology through the perspectives of higher education professionals. *Teaching and Teacher Education*, 152, 104774. <https://doi.org/10.1016/j.tate.2024.104774>
- Luke, A., & Freebody, P. (2021). *Literacy as social practice: The new literacy studies*. Routledge.
- Martin, J. R., & Rose, D. (2022). *Working with discourse: Meaning beyond the clause* (2nd ed.). Bloomsbury.
- Maton, K., & Doran, Y. (2021). *Semantic waves and knowledge-building in language education*. Routledge.
- Mochizuki, N., & Starfield, S. (2021). Dialogic interactions and voice negotiations in thesis writing groups: An activity systems analysis of oral feedback exchanges. *Journal of English for Academic Purposes*, 50, 100956. <https://doi.org/10.1016/j.jeap.2020.100956>
- Moradian, M. R., Miri, M., & Alamdar, P. (2021). The role of dialogic interaction in EFL writing assessment: A sociocultural perspective. *Indonesian Journal of Applied Linguistics*, 10(3). <https://doi.org/10.17509/ijal.v10i3.31754>
- Morton, T. (2023). Afterword: SFL, theoretical pluralism and content and language integration at the levels of curriculum, pedagogy and assessment. *International Journal of Bilingual Education and Bilingualism*, 26(3), 363–368. <https://doi.org/10.1080/13670050.2020.1806781>
- Nadia, H., Yansyah, Y., & Rafidiyah, D. (2021). Designing an online work-related language skill for nursing students: A developmental research. *JEES (Journal of English Educators Society)*, 6(1), 163–170. <https://doi.org/10.21070/jees.v6i1.1112>
- Nimasari, E. P., Setiawan, S., Munir, A., & Suhartono, S. (2023). Why didn't I make it digital: Indonesian teachers' perception of dialogical-based online teaching and socio-emotional connections. *The Qualitative Report*, 28(8). <https://doi.org/10.46743/2160-3715/2023.6105>
- Nimasari, E. P., Setiawan, S., & Munir, A. (2024). How do Indonesian student-teachers experience wellbeing during research supervision? A qualitative interview study. *International Journal of Learning, Teaching and Educational Research*, 23(3), 348–366. <https://doi.org/10.26803/ijlter.23.3.17>
- Pakir, A. (2019). Professor Michael Alexander Kirkwood (M.A.K.) Halliday, 1925–2018, founding father of Systemic Functional Linguistics. *Asian Englishes*, 21(3), 324–328. <https://doi.org/10.1080/13488678.2019.1608720>
- Pass, S. (2007). When constructivists Jean Piaget and Lev Vygotsky were pedagogical collaborators: A viewpoint from a study of their communications. *Journal of Constructivist Psychology*, 20(3), 277–282. <https://doi.org/10.1080/10720530701347944>
- Ryshina-Pankova, M., Barthold, W., & Barthold, E. (2021). Enhancing the content- and language-integrated multiple literacies framework: Systemic functional linguistics for teaching regional diversity. *System*, 96, 102403. <https://doi.org/10.1016/j.system.2020.102403>
- Silseth, K., & Gilje, Ø. (2019). Multimodal composition and assessment: A sociocultural perspective. *Assessment in Education: Principles, Policy & Practice*, 26(1), 26–42. <https://doi.org/10.1080/0969594X.2017.1297292>
- Suwannasom, W., & Pattanajak, N. (2023). Developing an integrated model of digital learning ecosystem in higher education: A study in Thai universities. *Education and Information Technologies*, 28, 1535–1554. <https://doi.org/10.1007/s10639-022-11178-6>
- Thomas, D. P. (2022). Structuring written arguments in primary and secondary school: A systemic functional linguistics perspective. *Linguistics and Education*, 72, 101120. <https://doi.org/10.1016/j.linged.2022.101120>
- Van Der Vegt, I., Kleinberg, B., & Gill, P. (2023). Proceed with caution: On the use of computational linguistics in threat assessment. *Journal of Policing, Intelligence and Counter Terrorism*, 18(2), 231–239. <https://doi.org/10.1080/18335330.2023.2165137>
- Weinmann, M., Neilsen, R., & Cabezas Benalcázar, C. (2024). Linguaging and language awareness in the global age 2020–2023: Digital engagement and practice in language teaching and learning in (post-) pandemic times. *Language Awareness*, 33(2), 347–364. <https://doi.org/10.1080/09658416.2023.2236025>
- Wijayati, R. D., & Khafidhoh, K. (2023). Developing Islamic-values-based PowToon learning media for Muhammadiyah kindergarten teachers: COVID-19 and TEYL. *English Language Teaching Educational Journal*, 5(2), 103–118. <https://doi.org/10.12928/eltej.v5i2.6475>
- Zhang, L., & Wang, J. (2024). “Curriculum-based or ideology-based”: An action research on China's curriculum-based ideology and virtuous awareness education in second language writing teaching: An activity theory perspective. *Heliyon*, 10(11), e32631. <https://doi.org/10.1016/j.heliyon.2024.e32631>

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright ©2025 Elok Putri Nimasari, Bambang Harmanto, Rinjani Bonavidi, Adi Fajaryanto Cobantoro, Elvina Dwi Cahyani, Velrinda Fahza'a Risty. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Enhancing vocabulary acquisition through word association recitation among EFL students

**Andika Rikarno Putra,¹ Dodi Mulyadi,^{*1} Siti Aimah,¹ Pirman Ginting,²
Charanjit Kaur Swaran Singh,³**

¹Universitas Muhammadiyah Semarang, Indonesia, ²Universitas Muhammadiyah Sumatera Utara, Indonesia,

³Universitas Pendidikan Sultan Idris, Indonesia

Vocabulary is one of the key components of mastering English. Schools provide an important environment where EFL students can begin learning the language. This study aimed to examine whether word association recitation could enhance vocabulary acquisition among beginner-level EFL students at an Islamic junior high school. A Classroom Action Research (CAR) design was employed to investigate the implementation of the word association recitation technique in improving vocabulary acquisition among seventh-grade students at Nurus Sunnah Islamic Junior High School. This CAR was conducted over two cycles between October and November 2024, with each cycle consisting of three instructional sessions. A purposive sampling method was used to select Class 7A (N= 143), which included students with varied levels of prior English exposure. The findings showed improve classroom engagement: attention increased from 46% to 67%, participation from 47% to 71%, and positive attitudes from 57% to 74%. Notably, post-test scores rose by 46.7%, from an average of 49.23 to 72.23, with substantial gains among lower-proficiency learners (up to 122% improvement). A paired-samples t-test revealed a statistically significant difference between pre- and post-intervention scores ($p= .000$), confirming the positive impact of word association recitation on vocabulary learning. Future research is recommended to explore the integration of communicative tasks, assess the approach across broader EFL contexts, and incorporate digital tools to further enhance vocabulary practice.

Keywords: EFL students, vocabulary acquisition, word association recitation, classroom action research

OPEN ACCESS

ISSN 2503 3492 (online)

*Correspondence:

Dodi Mulyadi

dodi@unimus.ac.id

Received: 17th June 2025

Accepted: 6th September 2025

Published: 16th October 2025

Citation:

Putra, A.R., Mulyadi, D., Aimah, S.,
Ginting, P., Singh, C.K.S. (2025).

Enhancing Vocabulary Acquisition
through Word Association Recitation
among EFL Students. *JEES (Journal of
English Educators Society)*, 10(2).
<https://doi.org/10.21070/jees.v10i2.1950>

INTRODUCTION

One of the crucial aspects of language acquisition is vocabulary development. A strong vocabulary foundation enables students to communicate more effectively, comprehend texts with greater ease, and express their ideas clearly. Mastery of vocabulary also allows learners to engage actively in both academic and social discourse (Marlinton et al., 2023). At the beginner level, students who are exposed to English for the first time are still adapting to the language (Daar, 2020). At this stage, the primary learning goals include understanding simple instructions, everyday expressions, and basic conversations. As learners progress, another important objective is to enhance their reading abilities by recognizing and comprehending frequently used words. This may involve reading short and simple texts such as signs, labels, and basic stories. Through these gradual steps, learners expand their vocabulary, which is vital for overall language development. Effective vocabulary instruction is therefore essential in helping students establish a strong lexical foundation and engage with the target language more proficiently.

To use English in daily communication, students must first master the most frequently used words, known as "lemmas," which serve as the core units of meaning in a language. [Webb \(2016\)](#) suggests that focusing on 800 to 1,000 of the most common lemmas enables learners to quickly develop a functional grasp of the language. However, preliminary screening conducted by the English teacher at the beginning of the school year revealed that the seventh-grade students at Nurus Sunnah Islamic Junior High School lacked knowledge of even the most basic vocabulary (see [Appendix 1](#)). In this test, students were asked to translate 40 random words (nouns and verbs) selected from the 1,000 most commonly used words in English.

Not all teachers are adequately prepared to address such challenges, which often results in frustration and decreased motivation among EFL students ([Irma, 2020](#); [Abdalla & Alameen, 2019](#)). This lack of preparedness may stem from limited training in managing diverse learning needs or handling classroom difficulties specific to language acquisition. To break this cycle, teachers should adopt practical techniques that foster engagement and accommodate individual learner differences.

[Ben-Jacob and Glazerman \(2021\)](#) note that the learning process involves multiple form of interaction, and they found that learning by doing (practice) is far more effective than passive listening. Applied to language learning, this suggests that traditional vocabulary instruction methods, such as translation and memorization, could be enhanced by integrating recitation and word association as strategies for improving vocabulary acquisition. Supporting this idea, [Tryana & Mahmud \(2023\)](#) strengthens found that learning new languages involves associating unfamiliar words with prior knowledge. Similarly, [Rashid et al \(2022\)](#) emphasize the importance of teaching vocabulary in context, highlighting the need for learners to understand not only word meanings but also collocations, multiple uses, and grammatical functions.

The word association and recitation technique involve the deliberately new vocabulary with familiar concepts, images, or personal experiences. This method encourages learners to create meaningful connections between new words and their existing knowledge or lived experiences ([Gunawan, 2020](#)). Furthermore, repeated exposure to vocabulary through systematic practice promote incremental gains in word retention and recall ([Gunawan, 2020](#)). By leveraging the combined benefits of association and recitation, students can strengthen their lexical retention and ultimately enhance overall language proficiency.

Previous studies have employed various methods and media to enhance students' vocabulary, including crossword puzzle, movies application such as Quizlet, the word association method, and progressive word increments. However, none have investigated the combination of word association and recitation within a classroom action research setting. This study therefore examines the effectiveness of using word association recitation techniques to enhance vocabulary acquisition among EFL students at a private junior high school in Semarang, Indonesia. Guided by this objective, the study addresses two research questions: a) How can word association recitation be implemented to enhance

vocabulary acquisition among EFL students? and b) To what extent can word association recitation enhance vocabulary acquisition among EFL students?

METHODS

This study was primarily quantitative in nature, employing a classroom action research design while also incorporating qualitative insights to enrich the findings. The study focused on gathering empirical data to evaluate the extent to which word association and recitation technique enhanced students' ability to recall and use new vocabulary. Quantitative data were obtained from assessments administered at the end of each cycle to measure student learning outcomes. To complement this, qualitative data were gathered through field observations conducted collaboratively by the researcher and the English teacher using an observation sheet. This study received approved from the school principal.

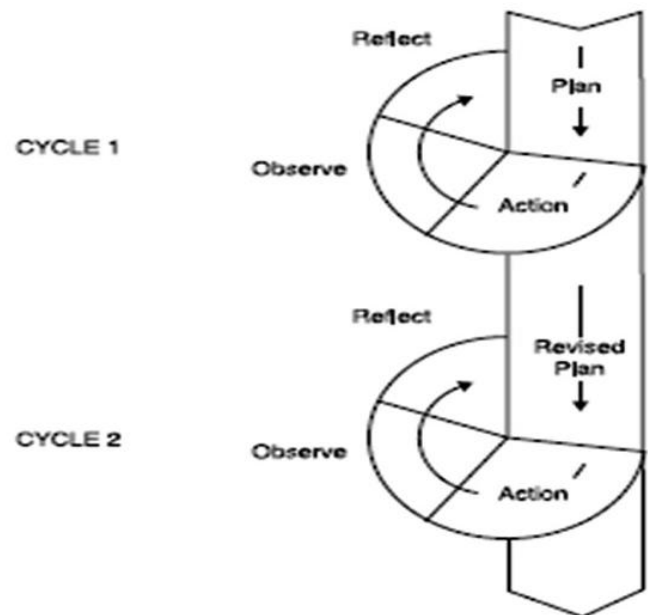


FIGURE 1 | Kemmis & Taggart's Classroom Action Research

This study followed the steps of action research outlined by [Kemmis and Taggart \(1988\)](#). The first step designing a general plan for action. In the action phase, the researcher implemented teaching based on the planned scenario, while another teacher conducted observation to document the effects of the intervention. The reflection phase entailed analyzing and evaluating these effects as the basis for further planning. At this phase, the outcome of the action was described and assessed. The general plan was then revised, and the subsequent action was planned, implemented, monitored, and evaluated.

This research was conducted in English class at a private Islamic junior high school in Semarang. The total student's population of the school was 143, consisting of 44 male and 99 female students. This study employed purposive sampling, selecting participants based on specific criteria. Grade seven was chosen because students at this level had just entered junior high school, making it appropriate to compare differences in skills between those with and without prior

formal English education. Class 7A, consisting of 13 students, was chosen because it included both students who had previously received English instruction in elementary school and those who had not.

Research Instruments

A research instrument is a tool or device used to collect, measure, and analyze data in a research study. The instruments used in this classroom action research were as follows:

Observation Sheets

The observation sheet in this study was developed based on indicators of the students' activities during the Word Association Recitation. These indicators included students' attention during learning activities, active participation in classroom tasks, and positive attitudes such as enjoyment or happiness toward the learning process. Further details about the observation sheet are provided in the Appendix.

The updated Vocabulary Levels Test

This study examined students' mastery of the 1,000 most frequently used words in English. To measure the vocabulary knowledge, it employed "The Updated Vocabulary Levels Test" developed by [Webb et al. \(2017\)](#). This test specifically measures receptive vocabulary knowledge, the ability to recognize the connection between a word's form and its meaning when reading. It does not measure productive vocabulary knowledge (the ability to produce a word) or other dimensions such as collocation or polysemy. To minimize score inflation from repeated testing, two equivalent forms (Form A and Form B) were used. The test evaluates vocabulary knowledge across five levels; however, this study focused on the section designed to measure mastery of the 1,000 most frequently used words.

Scores and Assessments

Scores and assessments were obtained at the end of each cycle to assess student learning outcomes. Students' scores were converted into percentages of the maximum possible score. To analyze the data, the researcher calculated the average vocabulary scores across cycles, providing an overall measure of students' vocabulary development.

Data Analysis

Data analysis is the process of systematically applying statistical and logical techniques to describe, summarize, and compare data. It involves interpreting the data collected during research to draw conclusions and address research questions or test hypotheses. In the context of classroom action research, the stages of sorting, describing, and inferring were used to interpret the interventions to guide subsequent classroom actions.

Data sorting was carried out using applications such as Microsoft Excel and SPSS version 26. Vocabulary test results were inserted into Excel to identify patterns, while daily activity data were processed in SPSS to conduct a paired-samples t-tests.

Descriptive analysis involved organizing and presenting results from each instrument. Vocabulary test results were presented in tables showing pre-test and post-test scores. Observation sheet data were presented in graphs, and daily assessment results were presented in tables.

Inferential analysis was conducted by comparing the results of the vocabulary tests with observations from cycles 1 and 2. Additionally, daily assessments collected throughout the CAR were analyzed against predetermined success indicators. Success indicators functioned as benchmarks for determining the effectiveness of the intervention. According to [Zain and Djamarah \(2006\)](#), the success of a learning process is achieved when at least 75% of students reach a minimum, optimal, or maximum level of success, allowing the learning process to progress to subsequent subjects. Based on this criterion, the success indicators for this present classroom action research were defined as follows:

Success Indicators for Students' Vocabulary Mastery: The CAR was considered successful if at least 80% of students demonstrate an adequate level of vocabulary mastery. Success was defined as students achieving a minimum score of 80% on the vocabulary mastery assessments, as measured against the established indicators.

Learning Outcome Success Indicator: The CAR was considered successful if at least 80% of students meet the minimum completeness criteria for learning outcomes at the end of each cycle. In this context, success means that the majority of students achieve the required level of understanding and performance as specified in the predetermined learning objectives.

RESULTS AND DISCUSSION

The researchers conducted the action research over three cycles, each consisting of the stages of planning, action, observation, and reflection. Each cycle comprised two lessons. The Grade 7A English class was scheduled every Tuesday during the semester. The first meeting, held on October 29, 2024, served as the introduction and administration of the pre-test. The following are details of each cycle.

Cycle 1

The learning material for Cycle 1 was the Chapter 2 "Daily Conversation." This cycle consisted of three lessons. The stages were as follows:

Lesson 1

Planning

The action plans for Cycle I included: (1) Preparing learning tools such as lesson plans and handouts; (2) Developing a minimum completeness criteria sheet and evaluation test questions; (3) Preparing observation sheets for monitoring student activities; (4) Preparing documentation tools.

Acting

During the lesson on October 5th, students engaged in various structured activities designed to improve their comprehension and application of to be in context. The main activities included listening to a recording of a conversation between two friends attending a concert, reading the corresponding dialogue, and determining whether given statements about the dialogue were true or false (Activity 2). Following this, the teacher explained the concept and usage of to be (is, am, are) with relevant examples. Students then worked in pairs to discuss related or similar words found in the dialogue. The

session concluded with Activities 3 and 4, which completed either individually or collaboratively in pairs.

Observing

During this lesson, students were unfocused and unable to answer questions when asked.

Reflecting

Based on the observation and activity results, the researcher concluded that the students' vocabulary mastery was still not satisfactory, as reflected in their lack of response. The teacher decided to familiarize the students with classroom instructions in English to increase their engagement and comprehension.

Lesson 2

Planning

Building on the findings from the first lesson, the second one was designed to include more oral practices, such as reading new words aloud before engaging with them in written form.

Acting

The second meeting, held on November 12, 2024, focused on the theme of nationalities and the vocabulary associated with them. This session aimed to expand students' understanding of countries and the corresponding terms used to describe their citizens. The main activities included: 1) listening to a recording that introduced the pronunciation of various country names in English (5 minutes). 2) Participating in an oral repetition exercise to practice pronunciation (5 minutes). 3) Reading a short text about Cappadocia from the textbook (p. 27) to contextualize the vocabulary (10 minutes). 4) Studying a list of country names alongside their corresponding nationalities (p. 29) to further build vocabulary (15 minutes).

Observing

During this lesson, students reacted slowly to new tasks. Those who did not understand the instructions tended to ask their peers for clarification before engaging in the activity.

Reflecting

Based on the observation and activity results, the researcher concluded that the students' vocabulary showed improvement, although it was still not fully satisfactory.

Lesson 3

Planning

This lesson took place immediately after the previous one, with only a lunch break in between. According to the schedule, this class had two consecutive lessons once every week. This arrangement limited opportunities for advance reparation, so the teacher still continued using the previous strategy of familiarizing students with classroom instructions in English.

Acting

The third meeting, held on November 12th, 2024, focused on nationalities and related vocabulary, with a specific focus on forming and using WH-questions in context. The main activities included: 1) explaining WH-question words, who, what, where, and how, and their correct usage in sentences (15 minutes). 2) Completing Activity 9, which required students to apply WH-questions appropriately (15 minutes). 3) Working on Activities 10 and 11, where students completed dialogues by inserting the correct WH-questions

based on visual prompts (25 minutes). 4) Engaging in a closing reflection, reporting on the progress of their mini projects, and receiving information about the upcoming session (10 minutes).

Observing

During this lesson, students were mostly focused and had begun answering questions that they understood.

Reflecting

Based on the observation and activity results, the researcher concluded that students' vocabulary mastery was still below expectations. This was evident from the number of students who scored below the minimum mastery criterion. The researcher decided to provide more extensive vocabulary practice along with additional pronunciation exercises in the following lessons.

The data in [Table 1, appendix 1](#) presents the scores for 13 students across three lessons (Lesson 1, Lesson 2, and Lesson 3) in Cycle 1. Overall, Lesson 2 showed the highest performance, with a mean score of 78.77 and a median of 80, while Lesson 3 recorded the lowest mean (74.46) and median (74), indicating it was the most challenging session. Student A consistently obtained lower scores, with an overall mean of 71.33, whereas, student AFA performed strongly across all lessons, achieving the highest mean score (83.00).

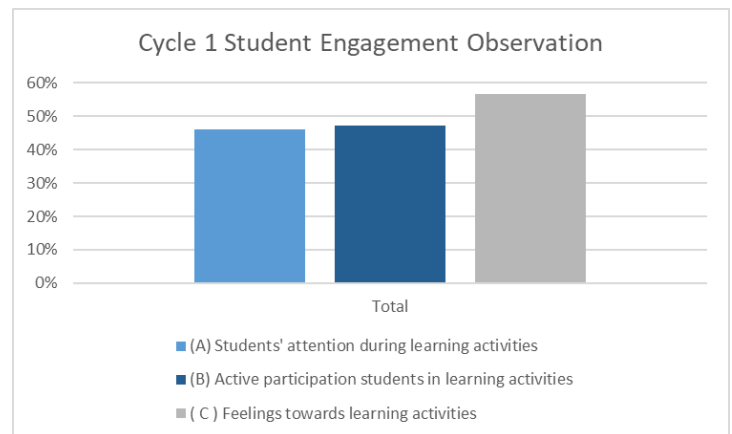


FIGURE 1 | Cycle 1 Student Engagement Observation

[Figure 1](#) describes the student engagement as observed in the first cycle. Students demonstrated the highest engagement in terms of positive attitudes and emotional investment (57%). However, their focus during lessons (46%) and active involvement in tasks (47%) were slightly lower, suggesting a need for improvement in sustaining concentration and participatory behaviors. Overall, the results indicate a moderate level of engagement, with emotional factors (C) surpassing both cognitive and behavioral aspects (A and B).

Cycle 2

The learning material for Cycle II was a continuation of Cycle I, with a specific focus on "Nationality." The cycle aimed to refine and enhance the effectiveness of the teaching strategies based on feedback and data collected from the previous cycle.

Lesson 1

Planning

The action plans for Cycle II included: (1) preparing learning tools such as lesson plans and handouts; (2) developing a

minimum completeness criteria sheet and evaluation test questions; (3) preparing implementation of word association games to all students; and (5) preparing documentation tools.

Acting

The second cycle was conducted on November 19th, 2024. The session began with a review activity where students were asked reflective questions from the previous meeting, such as: "How do you use 'what, who, where, and how' in sentences?" This served as a warm-up to activate prior knowledge. During the main activities (55 minutes), the teacher introduced a Word Association Game designed to strengthen vocabulary retention and reinforce the contextual use of the verb to be. Students were first encouraged to pair up and discuss similar or related words they had encountered in earlier prior lessons (5 minutes). Following this, they were divided into small groups of 2 - 4 members and tasked with creating a short dialogue that included self-introduction and the use of WH-questions to gather information (25 minutes). Each group then performed their dialogues in front of the class, giving them the opportunity to apply their learning in a communicative and interactive manner (25 minutes).

Observing

Students demonstrated greater confidence because they had the chance to discuss with their partners before expressing their own opinions. The word association game also created a livelier classroom atmosphere, as evidenced by the active participation of all students in the activity.

Reflecting

The researcher noted a modest improvement in students' engagement during the lesson. It was decided to incorporate word association recitation in the following lesson to further enhance vocabulary retention.

Lesson 2

Planning

This lesson was designed to maximize vocabulary acquisition by (1) introducing including new vocabulary through a video (2) engaging students in group tasks to encourage discussions of the new vocabulary; and (3) reinforcing memorization through word association recitation.

Acting

The second meeting, held on November 23rd, 2024, focused on teaching students about nationalities and related vocabulary. This session aimed to enhance students' cultural awareness and descriptive abilities through engaging and interactive activities. During the core session (55 minutes), students were first asked to form groups of three (5 minutes) and re-watch their selected video featuring a famous place (15 minutes). After viewing, they noted key information, including the location, the reason the place is famous, and their personal reasons for liking it (15 minutes). To consolidate vocabulary and memory skills, the teacher then introduced a Word Association Recitation activity (20 minutes). In this task, the teacher wrote a base word on the board (e.g., sun), and students took turns suggesting associated words (e.g., round, circle, wheel, car, etc.) until approximately eight words were listed. Each student silently chose one word to memorize. The teacher then confirmed word selections by asking for a show of hands and ensured that any unchosen words were memorized by the entire class.

All words were then erased and replaced with circles. As the teacher pointed to each circle, students recalled and pronounced their chosen word aloud. For unselected words, the whole class either clapped in unison or recited the word together. The process was repeated forwards and backwards at varying speeds, concluding with a full-class recitation of all the words. This activity not only strengthened vocabulary development but also fostered engagement, memory retention, and collaborative learning.

Observing

Students demonstrated improved focus and stronger engagement in group discussions and task completion. While a few students still showed occasional restlessness, the overall learning atmosphere was noticeably more productive.

Reflecting

The researcher observed clear progress in the students' vocabulary and pronunciation skills.

Lesson 3

Planning

This lesson was conducted immediately after the previous one, following a lunch break. According to the schedule, this class had two consecutive sessions once every two weeks. This limited the planning process; therefore, the teacher continued the previous strategy of strengthening student's vocabulary acquisition, followed by a daily assessment.

Acting

The third meeting, held on November 23rd, 2024, began with students asking questions about any material from previous sessions that they had not yet fully understood. As a warm-up activity, the class engaged in a Word Association Recitation (15 minutes) to stimulate memory and language fluency. In this activity, the teacher wrote a word on the board, and students collaboratively generated associated words. The sequence was repeated forwards and backwards at varying speeds to strengthen retention, concluding with the class reciting all the words in unison. Following this activity, students completed the final assignment for the chapter (Textbook pp. 34–36) as part of the chapter evaluation. This task served as a summative assessment to measure their understanding of the material covered throughout the unit. Afterwards, students reflected on the second chapter by responding to several guiding questions posed by the teacher. The session concluded with a post-test on new vocabulary.

Observing

Students demonstrated better focus and stronger engagement in group discussions and task completion. They were notably more active in asking and answering questions. Although a few students still showed occasional restlessness, the overall classroom atmosphere was markedly more productive.

Reflecting

The researcher observed a modest improvement in students' vocabulary mastery, as reflected in the increased number of students who scored above the minimum passing grade (KKM). Additionally, noticeable progress was also evident in their pronunciation skills.

The data in [Table 2 Appendix 1](#) presents the performance of 13 students across three lessons (Lesson 1, Lesson 2, and Lesson 3) during Cycle 2. Overall, Lesson 1 recorded the

highest mean score (84.23), with a median of 90 and a mode of 90 (appearing 8 times), indicating strong and consistent performance. Lesson 3 followed closely, with a mean of 87.23, a median of 88, and two mode - 84 and 90 (each appearing 4 times). Lesson 2 showed a slightly lower mean (82.38), a median of 80, and a mode of 80 (5 occurrences), suggesting it was the most challenging of the three lessons. At the individual level, Student FAJ stood out with the highest score in Lesson 2 (96) and a strong overall mean (91.33). Student MZA achieved the single highest score (100 in Lesson 3), contributing to their high overall mean (90.00). In contrast, Student A obtained the lowest mean (76.67), with consistent but modest performance scores across all lessons.

Cycle 2 demonstrated a general improvement compared to Cycle 1. The mean scores rose significantly across all lessons, with Lesson 1 rising from 76.46 to 84.23 and Lesson 3 from 74.46 to 87.23. The mode in Lesson 1 also shifted markedly, from 74 in Cycle 1 to 90 in Cycle 2, achieved by 8 out of the 13 students, indicating that the use of word association recitation contributed to higher achievement. Even the lowest scores improved, with no student scoring below 75 in Cycle 2, compared to a minimum score of 66 in Cycle 1.

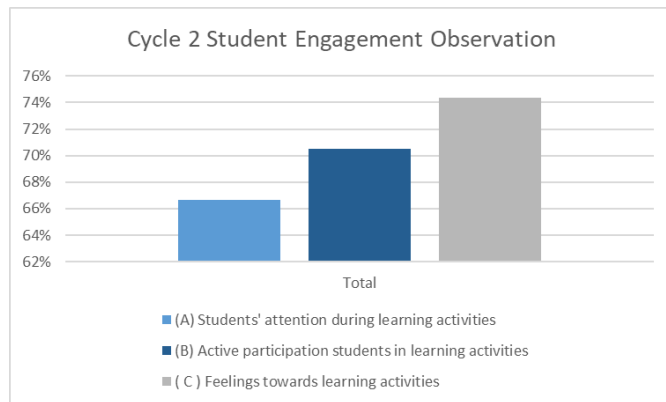


FIGURE 2 | Cycle 2 Student Engagement Observation

Figure 2 demonstrates significant improvements in student engagement during Cycle 2 compared to Cycle 1. Attention during learning activities (A) increased to 67% (from 46% in Cycle 1), active participation (B) increased to 71% (from 47%), and positive feelings toward learning (C) reached 74% (from 57%). These notable gains suggest that instructional adjustments made between cycles, such as incorporating more interactive word games and collaborative activities, effectively enhanced both behavioral and emotional engagement. The highest score in category (C) reinforces the critical role of motivation and classroom climate in language learning, while the balanced growth across all three indicators (A, B, C) reflects a more holistic improvement in student involvement.

To determine whether there was a significant difference between Cycle 1 and Cycle 2, a comparative test was employed. To select the appropriate test, the researcher first examined the data distribution using the Shapiro–Wilk normality test. The average scores from each cycle (Lesson 1, Lesson 2, and Lesson 3) were calculated and tested for normality. The significance value (Sig.) for the average of Cycle 1 was 0.580, and for Cycle 2 it was 0.507.

Since both values were greater than 0.05, it was concluded that the data from both cycles did not significantly deviate from a normal distribution, thus meeting the assumption of normality.

TABLE 3 | Paired T Test Results of Cycle 1 and Cycle 2

Name	Average		Paired T Test
	Cycle 1	Cycle 2	
A	71,33	76,67	Mean: -8.66718 Standard Deviation: 3.38335 Std Error Mean: 0.93837 t-value: -9.236 df: 12 P (2-tailed): 0.000
AFA	83,00	88,00	
AAA	72,67	85,00	
BFA	77,33	84,67	
F	74,00	79,67	
FAJ	81,00	91,33	
GBZK	79,00	84,67	
MI	72,67	80,67	
M	78,67	84,67	
MSD	76,33	84,33	
MZA	74,67	90,00	
NMX	74,67	85,00	
PAR	74,67	88,00	

Given the normal distribution of the data, a paired-samples t-test was conducted to examine whether the mean scores of the two cycles differed significantly. Table 3 presents the results. The paired t-test revealed a statistically significant difference between the two cycles, with a mean difference of -8.67 (SD = 3.38, SE = 0.94). The obtained t-value was -9.236 with 12 degrees of freedom, and the p-value was .000 (p < .05), indicating a highly significant difference. These findings show that the mean scores in Cycle 2 were significantly higher than those in Cycle 1, suggesting that the implementation of the word association recitation technique positively impacted students' vocabulary acquisition.

The next instrument was “The New Vocabulary Test” administered before and after the CAR. A comparison of the Pre-Test and Post-Test scores revealed significant progress in learners' vocabulary acquisition following the completion of both instructional cycles. The average Pre-Test score across all students was 49.23, which increased to 72.23 in the Post-Test, a remarkable improvement of 23 points (46.7%). Notably, every student showed progress, with the most dramatic gains observed among those who began with lower proficiency. For instance, Student F improved from 27 to 60 (a 122% increase), and Student AAA rose from 47 to 73 (55.3%). Even high-performing students, such as NMX, who started with a score of 87, advanced to 93, demonstrating that the instructional method was beneficial for learners at all proficiency levels.

The results underscore the overall effectiveness of the word association recitation. The most outstanding progress was seen in Student FAJ, who improved from 43 to 90 (109.3%), and Student AFA, who rose from 60 to 80 (33.3%). The consistent upward trend across all learners, regardless of their initial proficiency or prior experience, indicates that the methodology effectively addressed diverse learning needs.

These findings provide strong evidence for the value of the implemented teaching strategies in enhancing vocabulary acquisition among EFL students.

TABLE 4 | Pre-Test and Post-Test Score

Name	English in Primary School	Pre-Test	Post Test
A	No	37	50
AFA	Yes	60	80
AAA	Yes	47	73
BFA	No	27	53
F	No	27	60
FAJ	Yes	43	90
GBZK	Yes	73	93
MI	Yes	60	77
M	Yes	53	83
MSD	Yes	47	73
MZA	No	30	53
NMX	No	87	93
PAR	Yes	53	83

Learning English as a Foreign Language (EFL) presents significant challenges for beginners, particularly in the initial stages. The study’s findings underscore the tangible impact of word association recitation as an effective intervention for EFL learners. Prior to its implementation, observational data revealed noticeable gaps in vocabulary mastery between students with prior English education and those without. These learners struggled not only to retain new words but also to pronounce the accurately, reflecting a common challenge in early-stage language acquisition, where fragmented knowledge impedes both comprehension and expression (Nation, 2001).

Without a foundational lexicon, students struggle to comprehend instructions, express ideas, or engage in simple conversations, which can lead to frustration and disengagement (Krashen, 1982). The psychological impacts of these difficulties often manifest as anxiety and self-doubt, further raising the affective filter, a mental block that impedes language processing (Kiruthiga & Christopher, 2022).

Vocabulary acquisition is an ongoing process of learning new words for everyday use serves as the foundation for mastering any language (Baker, 2015). It involves helping students to understand the meanings of new words and concepts across various contexts and academic content areas. Building on the principles of gamified learning, word games, such as crosswords, Scrabble, word searches, or digital platforms like Wordle, offer targeted benefits for EFL learners. These activities align with the testing effect, a cognitive phenomenon in which the active retrieval of information strengthens long-term memory (Kiruthiga & Christopher, 2022). Supporting this, a study by Kabiri and Ghafoori (2015) demonstrated that word games not only improved immediate vocabulary recall but also enhanced long-term retention, as shown by delayed post-test scores

administered two weeks after treatment.

The Implementation of Word Association Recitation with EFL Students

The CAR is considered successful if at least 80% of students meet the minimum completeness criteria for learning outcomes at the end of each cycle. This indicates that the majority of students should achieve the required level of understanding and performance as outlined in the learning objectives. Student achievement is primarily measured through their performance on assessments, which are designed not only to reinforce understanding but also to evaluate learning outcomes. In Indonesia, the recommended minimum passing grade for English at the junior high school level is 75.

During the first cycle, the students’ average score was 74.5, slightly below the minimum passing grade. More than 60 % of students received failing scores, with the lowest being 66. Many Indonesian junior high school students struggle with learning English due to factors such as ineffective teaching methods, lack of motivation, and limited practice opportunities (Lestari, 2023). In addition, some students experience anxiety when learning a foreign language, which creates barriers to interaction and inhibits their ability to express themselves verbally (Hidayati, 2022).

In the second cycle, the class median increased to 82, with all students scoring above the minimum passing grade. This improvement may be attributed to the cycle’s focus on reading and pronouncing the text collectively. When students read aloud together after hearing the teacher’s model, their confidence increases as they practice pronunciation collaboratively (Wahyuni, 2022).

To determine whether there was a significant difference between Cycle 1 and Cycle 2, a paired t-test was used. The analysis revealed a statistically significant difference between the two cycles. The average scores in Cycle 2 were higher than those in Cycle 1, indicating that the implementation of the word association recitation technique positively impacted students’ vocabulary acquisition Cycle 2.

The Use of Word Association Recitation in Enhancing Vocabulary Acquisition among EFL Students

Success in this study was defined by achieving a minimum of 80% in assessments related to vocabulary mastery, based on the established measurement indicators. The assessment tool used to measure vocabulary mastery was “The New Vocabulary Level Test”.

The comparison between Pre-Test and Post-Test scores revealed notable progress in learners' vocabulary acquisition after completing both instructional cycles. The average Pre-Test score was 49.23, which increased to 72.23 in the Post-Test, a remarkable improvement of 23 points (46.7%). Notably, every student demonstrated progress, with the most dramatic gains observed among those who began with lower proficiency. These results underscore that word association recitation effectively enhanced vocabulary acquisition. The consistent upward trend across all learners, regardless of their starting level or prior experience, indicates that the methodology successfully addressed diverse learning needs.

Post-intervention results further demonstrated measurable, albeit modest, improvement. The number of

students achieving scores above the KKM increased, and a statistically significant difference was found between the two cycles, suggesting enhanced vocabulary retention. These findings align with Hasyim's (2018) study, which reported statistically significant improvement in the vocabulary at 31 grade VIIIA students at SMP Muhammadiyah Al Amin Sorong through the use of word association games. Similarly, Lube & Nuraeni's (2020) found significance vocabulary gains among 30 seventh-grade students at SMP Muslimin Cililin through word games. In line with these, Totti (2024) conducted classroom action research with grade VIII C students at SMP Santa Monika, reporting a 70% improvement in completion rates between cycles.

Limitations and Implications

This study has several limitations that should be acknowledged. First, because it was conducted as classroom action research, the teacher's dual role as both instructor and evaluator may have introduced bias in the implementation and assessment processes. Second, the duration of the study was restricted to two cycles, as data collection needed to be completed before the semester examinations (October 29th - December 2nd, 2024). Third, challenges arose in the ethical procedures, as only the principle's written agreement and parents' verbal consent were obtained due to time constraints. Fourth, the observational sheet employed had not undergone prior reliability and validity testing, which may have affected the accuracy of the classroom data collected.

Additionally, the research was conducted in a single private Islamic junior high school in Semarang, which may limit the generalizability of findings to other EFL contexts with different cultural or institutional settings. Despite these constraints, the results demonstrate the potential of word association recitation as a practical and accessible strategy for improving vocabulary acquisition. Future studies are encouraged to obtain comprehensive written consent, employ validated observation instruments, extend the duration of intervention cycles, and include a broader range of schools to enhance the reliability, validity, and applicability of the findings.

CONCLUSION

The study demonstrated that the word association recitation technique effectively enhanced vocabulary acquisition among EFL students. Quantitative results from the paired samples t-test revealed a statistically significant improvement in vocabulary scores from Cycle 1 to Cycle 2, with a mean difference of -8.67, a t-value of -9.236, and a p-value of .000, indicating substantial progress. Additionally, results from the New Vocabulary Test conducted before and after the intervention confirmed its improvement.

Integrating word association recitation into classroom routines can serve as a low-cost and accessible strategy for addressing vocabulary gaps, particularly in resource-limited contexts. Incorporating interactive, gamified techniques into junior high school English programs is recommended to enhance both learning outcomes and student motivation.

ACKNOWLEDGEMENTS

The authors would like to express their sincere appreciation to all participants, colleagues, and research assistants who contributed to the completion of this study. Special thanks are extended to the English Education Department of Universitas Muhammadiyah Semarang for the academic and institutional support that made this research possible.

REFERENCES

- Abdalla, Amir, & Alameen, Amir Minalla (2019). Factors Affecting Sudanese EFL Teachers' Teaching Performance at Public Schools Comparing to Their Performance at Private Schools.
- Baker, C. E. (2015). Fathers' language input during shared book activities: Links to children's kindergarten achievement. *Journal of Applied Developmental Psychology*, 36, 53–59. <https://doi.org/10.1016/j.appdev.2014.11.009>
- Ben-Jacob, M., & Glazerman, A. (2021). Technology and Education: A Merger with the Past, Present, and Future. *Open Journal of Social Sciences*, 09, 39–42. <https://doi.org/10.4236/jss.2021.94004>
- Daar, G. F. (2020). Problems Of English Language Learning in Context (Based on some Studies in Manggarai). *JPKMI (Jurnal Pengabdian Kepada Masyarakat Indonesia)*, 1(2), 116–124.
- Gunawan, H. (2020). The Effectiveness of Repetition in Increasing Students Vocabulary. *Simetris*, 14(2), 49–52.
- Hasyim, R. (2018). Developing Vocabulary Mastery Through Word Association by The Second Year Students of SMP Muhammadiyah. *Jurnal Qalam: Jurnal Ilmu Kependidikan*, 1–5. https://www.academia.edu/39270912/developing_vocabulary_mastery_through_word_association_by_the_second_year_students_of_smp_muhammadiyah
- Hidayati, A. N., Abdullah, F., Andriani, A., Rosmala, D., & Nurvianti, N. (2022). English Speaking Anxiety Among Indonesian Junior High School Learners: in Search of Causes and Solutions. *Getsempena English Education Journal*, 9(1), 53–63. <https://doi.org/10.46244/geej.v9i1.1746>
- Irma. (2020). Challenges Encountered by EFL Teachers in Teaching Productive and Receptive Skills at SMP 2 Mattirobulu Pinrang. *Journal of Islamic Studies and Society*, 1(2), 133–158.
- Kabiri, R., & Ghafoori, N. (2015). The impact of playing word games on young Iranian EFL learners' vocabulary learning and retention. *The Journal of Applied Linguistics*, 7(14), 67–82.
- Kiruthiga, E., & Christopher, G. (2022). The Impact of Affective Factors in English Speaking Skills. *Theory and Practice in Language Studies*, 12(12), 2478–2485. <https://doi.org/10.17507/tppls.1212.02>
- Kemmis, S., & McTaggart, R. (1988). *The Action Research Planner*. Victoria: Deakin University Press.

- Krashen, S. D. (1982). *Principles and Practice in Second Language Acquisition (First)*. Pergamon Press Inc.
- Lestari, P. I. (2023). Exploring Students' Attribution: Why Did Students Fail in Their Online Basic Grammar Class? *Journal of English and Education (JEE)*, 9(1), 22–32. <https://doi.org/10.20885/jee.v9i1.28046>
- Lube, C. I., & Nuraeni, S. (2020). Improving English Vocabulary Mastery Through Word Game. *PROJECT (Professional Journal of English Education)*, 3(1), 109. <https://doi.org/10.22460/project.v3i1.p109-113>
- Marlinton, M., Syahri, I., & Mayasari, S. (2023). The Efficacy of Task-Based Learning and Learning Motivation on Student's Reading Comprehension: A Wholistic Study. *Elsya: Journal of English Language Studies*, 5(2), 250–269. <https://doi.org/10.31849/elsya.v5i2.14082>
- Minalla Alameen, A. (2022). From EFL Teachers' Perspective: Impact of EFL Learners' Demotivation on Interactive Learning Situations at EFL Classroom Contexts. *English Language Teaching*, 15, 1. <https://doi.org/10.5539/elt.v15n3p1>
- Nation, I. S. P., & Newton, J. (2008). *Teaching ESL/EFL Listening and Speaking*. Routledge. (<https://doi.org/10.4324/9780203891704>)
- Rashid, M. H., Lan, Y., & Hui, W. (2022). The importance of vocabulary in teaching and learning in applied linguistics. *Linguistics and Culture Review*, 6(March), 1548–1557. <https://doi.org/10.21744/lingcure.v6ns2.2177>
- Totti, C., Kurniawati, T., Darajat, A., & Wathoni, H. (2024). Improving Students' Vocabulary Learning Using Finding Word Game: A Classroom Action Research. *Journal of English Language Proficiency*, 1(1), 25-37. <https://doi.org/10.36312/jelap.v1i1.3872>
- Tryana, & Mahmud, L. H. (2023). Vocabulary Teaching by Words Association. *JiIP (Jurnal Ilmiah Ilmu Pendidikan)*, 6(9), 7246–7253. (<http://jiip.stkipyapisdmpu.ac.id>)
- Wahyuni, P. (2022). The Effect of Teaching Reading Aloud Technique in Improving English Pronunciation. *EBONY: Journal of English Language Teaching, Linguistics, and Literature*, 2(1), 46–54. <https://doi.org/10.37304/ebony.v2i1.4041>
- Webb, S., Sasao, Y., & Ballance, O. (2017). The updated Vocabulary Levels Test. *ITL - International Journal of Applied Linguistics*, 168(1), 33–69. <https://doi.org/10.1075/itl.168.1.02web>
- Zain, A. & Djamarah, S.B. (2006). *Strategi Belajar Mengajar*. Jakarta: Rieneka Cipta.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright ©2025 Andika Rikarno Putra, Dodi Mulyadi, Siti Aimah, Pirman Ginting, Charanjit Kaur Swaran Singh. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Appendix I
Score of Cycle 1 and 2 Learning Assessments

TABLE 1 | Score of the First Cycle

Name	Cycle 1		
	Lesson 1	Lesson 2	Lesson 3
A	70	74	70
AFA	84	85	80
AAA	72	80	66
BFA	74	74	84
F	80	72	70
FAJ	80	85	78
GBZK	78	85	74
MI	74	72	72
M	76	80	80
MSD	74	85	70
MZA	76	76	72
NMX	70	80	74
PAR	70	76	78

TABLE 2 | Score of the second Cycle

Name	Cycle 2		
	Lesson 1	Lesson 2	Lesson 3
A	75	75	80
AFA	90	84	90
AAA	90	75	90
BFA	90	80	84
F	75	80	84
FAJ	90	96	88
GBZK	90	80	84
MI	75	83	84
M	90	80	84
MSD	75	90	88
MZA	90	80	100
NMX	75	90	90
PAR	90	80	94

Appendix II
Observation Sheets

Observation sheets are records containing the results of observations conducted collaboratively by the researcher and the teacher, focusing on the variables under study. In this research, the observation sheet was designed based on indicators of activities carried out during the Word Association Recitation. These included students’ attention during learning activities, their active participation in classroom tasks, and their positive feelings toward the learning process.

The indicators for the observation sheets in this classroom action research centered on three key criteria: students’ attention, active participation, and feelings toward learning activities. Attention was measured by students’ ability to remain focused on the teacher’s explanations, tasks, and discussions, including maintaining eye contact, listening attentively, and responding when prompted. Active participation was evaluated through students’ engagement in classroom activities such as asking questions, contributing to discussions, and completing individual or group tasks, reflecting their involvement and initiative in the learning process. Lastly, feelings toward learning activities were assessed by observing students’ body language, facial expressions, and verbal feedback to gauge enthusiasm, motivation, or discomfort during lessons.

Together, these indicators provided a comprehensive picture of students’ engagement and emotional responses, which are essential for evaluating the effectiveness of the teaching strategies employed in the classroom. The criteria for assessing students’ vocabulary mastery through word association recitation were calculated using the following formula:

Formula	Value Range	Criteria
Percentage = (Score obtained / Maximum score) x 100 %	$0\% \leq VR < 30\%$	Insufficient
	$30\% \leq VR < 60\%$	Adequate
	$60\% \leq VR \leq 100\%$	Good

Cycle 1 Observation Sheet Scores

Indicator	Descriptor	Score
(A) Students' attention during learning activities	1. Focusing on the teacher's explanation	6
	2. Taking notes on the teacher's explanation	5
	3. Reading the teaching material	7
	4. Participating seriously in the assigned tasks	6
	5. Showing eagerness to participate in the tasks	4
	6. Being persistent and not easily discouraged in completing tasks	8
(B) Active participation in learning activities	1. Answering the teacher's questions	8
	2. Asking the teacher questions	8
	3. Expressing opinions	6
	4. Actively making observations	4
	5. Actively participating in classroom activities	5
	6. Responding to classmates' answers	6
	7. Discussing with group members	7
	8. Summarizing learning outcomes	5
(C) Feelings towards learning activities	1. Coming to class on time	9
	2. Bringing stationery and books	11
	3. Having complete notes of teaching materials	6
	4. Completing assignments on time	8
	5. Taking responsibility for tasks	7
	6. Remaining calm and not restless during lessons	4
	7. Not easily bored and give up on tasks	4
	8. Being cheerful during classroom activities	10

Cycle 2 Observation Sheet Scores

Indicator	Descriptor	Cycle 2
(A) Students' attention during learning activities	1. Focusing on the teacher's explanation	9
	2. Taking notes on the teacher's explanation	8
	3. Reading the teaching material	10
	4. Participating seriously in the assigned tasks	9
	5. Showing eagerness to participate in the tasks	8
	6. Being persistent and not easily discouraged in completing tasks	8
(B) Active participation students in learning activities	1. Answering the teacher's questions	10
	2. Asking the teacher questions	7
	3. Expressing opinions	8
	4. Actively making observations	9
	5. Actively participating in classroom activities	10
	6. Responding to classmotes' answers	11
	7. Discussing with group members	12
	8. Summarizing learning outcomes	9
(C) Feelings towards learning activities	1. Coming to class on time	13
	2. Bringing stationery and books	13
	3. Having complete notes of teaching materials	8
	4. Completing assignments on time	10
	5. Taking responsibility for tasks	8
	6. Remaining calm and not restless during lessons	6
	7. Not easily bored and give up on tasks	8
	8. Being cheerful during classroom activities	12

Appendix III
The 1,000 Word Level Vocabulary Levels Test (Part A)

Name : _____

Date : _____

	boy	rent	report	size	station	thing
How big or small something is						
Place buses and trains go to						
Young man						

	ear	gold	lake	letter	office	people
Information sent to people						
Men and women						
Place for working						

	fellow	hat	ice	joke	light	system
Funny story						
Man or boy						
Something worn on your head						

	date	forest	mistake	news	record	shop
Latest information						
Place with many trees						
Something that is not right						

	bar	conversation	neighbor	rain	rubbish	shirt
Person who lives nearby						
Things that are thrown away						
Type of clothing						

	continue	cook	phone	pull	sail	share
Hold and move something toward yourself						
Keep happening						
Use together with others						

	enter	finish	happen	own	sing	worry
End						
Go inside						
Have something that is yours						

	arrive	collect	consider	glance	need	pack
Think about something						
Reach the place you are going						
Look quickly at something						

	affordable	beautiful	boring	dry	rough	tall
Higher than normal						
Not interesting						
Not flat						

	closed	dirty	empty	musical	orange	sad
Having nothing						
Not clean						
Unhappy						

Appendix IV
The 1,000 Word Level Vocabulary Levels Test (Part B)

Name : _____

Date : _____

	choice	computer	garden	photograph	price	week
Cost						
Picture						
Place where things grow outside						

	eye	father	night	van	voice	year
Body part that sees						
Parent who is a man						
Part of the day with no sun						

	center	note	state	tomorrow	uncle	winter
Brother of your mother or father						
Middle						
Short piece of writing						

	box	brother	horse	hour	house	plan
Family member						
Sixty minutes						
Way of doing things						

	animal	bath	crime	grass	law	shoulder
Green leaves that cover the ground						
Place to wash						
Top end of your arm						

	drink	educate	forget	laugh	prepare	suit
Get ready						
Make a happy sound						
Not remember						

	check	fight	return	tell	work	write
Do things to get money						
Go back again						
Make sure						

	bring	can	reply	stare	understand	wish
Say or write an answer to somebody						
Carry to another place						
Look at for a long time						

	alone	bad	cold	green	loud	main
Most important						
Not good						
Not hot						

	awful	definite	exciting	general	mad	sweet
Certain						
Usual						
Very bad						



From struggles to strategies: Indonesian students' use of e-portfolios in public speaking courses through the lens of self-regulated learning and sociocultural theory

Raden Muhammad Ali,^{*1} Pamastu Narpaduita,¹ Muhammad Iqwan Sanjani,²
Dwi Hastuti,¹ Ummi Khaerati Syam,³ Muhammad Kunta Biddinika¹

¹Universitas Ahmad Dahlan, Indonesia, ²University of New South Wales, Australia,

³Universitas Muhammadiyah Makassar, Indonesia

The implementation of e-portfolios in education has gained recognition as both an assessment and a learning tool. This descriptive qualitative study examines the challenges faced by students in creating electronic portfolios (e-portfolio) using Google Sites, consisting of speech drafts, recorded presentations, reflection notes, and peer feedback, and the strategies they employed to overcome these challenges. This study involved 22 English Education students enrolled in a public speaking course at a university in Yogyakarta, Indonesia, including eight participants who were interviewed. Data were collected through observations, semi-structured interviews, and documentation. The findings highlight three major challenges: technical barriers (digital skills and internet access), time management difficulties, and creativity-related issues. These findings are interpreted through the lenses of self-regulated learning theory and sociocultural theory, which explain how learners manage their own learning and benefit from peer and lecturer support. To address these challenges, students employed peer collaboration, self-directed learning strategies, and time management techniques. The study offers insights for designing e-portfolio-based speaking courses in similar EFL contexts by systematically addressing obstacles while fostering learner autonomy, creativity, and responsibility. Practically, it provides guidance for lecturers to scaffold e-portfolio tasks and integrate peer collaboration effectively. Methodologically, it demonstrates the application of thematic analysis in e-portfolio research. Theoretically, it links e-portfolio use with self-regulated learning and sociocultural theory.

Keywords: e-portfolio development, public speaking, student challenges, learning strategies, digital tool

OPEN ACCESS

ISSN 2503 3492 (online)

*Correspondence:

Raden Muhammad Ali

raden.ali@pbi.uad.ac.id

Received: 21th August 2025

Accepted: 14th August 2025

Published: 20th October 2025

Citation:

Ali, R.M., Narpaduita, P., Sanjani, M.I., Hastuti, D., Syam, U.K., Biddinika, M.K. (2025). From struggles to strategies: Indonesian students' use of e-portfolios in public speaking courses through the lens of self-regulated learning and sociocultural theory. *JEES (Journal of English Educators Society)*, 10(2). <https://doi.org/10.21070/jees.v10i2.1969>

INTRODUCTION

Portfolios were originally conceived as a way to showcase a selection of the best work for a specific audience. This artistic concept shifted during the 1980s, and portfolios are now widely used as educational tools to document students' progress, processes, competencies, and accomplishments over time (Bone et al., 2020). Paper portfolios gave way to electronic versions in the 1990s, while local networks were gradually replaced by the internet. The rapid growth of technologies like websites, blogs, Wikipedia, and Facebook made it feasible to develop multi-device e-portfolios, ensuring both efficacy and efficiency. In their most common

form, e-portfolios are digital adaptations of traditional paper portfolios (Scully et al., 2018). More broadly, e-portfolios can be defined as collections of web-based learning artifacts, such as documents, multimedia projects, speeches, dan images, accompanied by reflections that emphasize learning and its development (Ciesielkiewicz, 2019). The most widely accepted definition states that an e-portfolio is a collection of items and digital artifacts selected and managed by their owners (Isiyaku et al., 2018).

E-portfolios appear to be an effective tool for English education majors, supporting their learning across the four skills of speaking, writing, listening, and reading (Aulia et al., 2018). For instance, students can upload voice recordings and videos for listening classes, store reading materials for reading classes, preserve written assignments for writing classes, and record oral presentations in speaking classes. Many students experience anxiety in speaking classes due to limited practice opportunities, the complexity of the material, or shyness when responding in English. Audio and video recordings provide a promising strategy to increase their courage and self-confidence. Through e-portfolios, students also have the chance to create personal webpages, reflect on their learning, and document their process. They can monitor and assess their own development, which helps foster fluency in English. Additionally, lecturers can easily access student work via to e-portfolios integrated with information and communication technologies, provide timely feedback, and allow students to achieve this feedback for future reference.

Despite their potential, the implementation of e-portfolios in teaching and learning has proven to be more complex than initially anticipated. Although researchers and educators widely acknowledge the benefits of e-portfolios as both assessment and learning tools, numerous barriers hinder their successful adoption. Bolliger and Shepherd (2010) emphasize their potential, yet practical applications reveal persistent challenges. According to Zheng (2022), these include the need for robust technological infrastructure, sufficient institutional funding, and safeguards for user privacy. Additionally, varying levels of technological proficiency among educators and students complicate implementation. Overcoming these challenges require careful planning, adequate resource allocation, and targeted training to ensure the effective use. In the Indonesian context, students in public speaking courses often face difficulties in oral communication due to limited practice opportunities, fear of making mistakes, and low confidence. These challenges can negatively affect their learning outcomes and overall engagement.

E-portfolio systems may be robust and controllable, but their underlying design model can be constrained and inflexible in managing the presentation and flow of material, including text, multimedia, and web links. For institutions preparing to implement or update their e-portfolio systems, functionality often presents a challenge. One key issue is determining whether a well-established e-portfolio infrastructure can provide a supportive environment for students to engage in meaningful reflection that enhances learning quality. Nagle et al. (2019) noted that one of the main challenges in facilitating students' active engagement in the e-portfolio process is fostering intrinsic motivation. Both self-

directed learning and the adoption of new technologies rely heavily on intrinsic motivation. Unlike extrinsic motivation, which is driven by external rewards, pressures, or incentives, intrinsic motivation stems from enjoyment, interest, or a sense of challenge (Ciesielkiewicz, 2019). Despite the growing global literature on e-portfolios, few studies have examined how Indonesian EFL students experience and respond to challenges in creating e-portfolios for public speaking courses. Previous research has primarily focused either on the general benefits of e-portfolios in language learning or on technical aspects of implementation, but has rarely addressed the combined challenges of technical barriers, time management, and creativity-related struggles within this specific context. This gap highlights the urgency of investigating the phenomenon, particularly as e-portfolio adoption in Indonesian higher education continues to expand without sufficient empirical evidence to guide effective implementation. Accordingly, this study aims to investigate the challenges faced by Indonesian students when developing e-portfolios in public speaking courses and to explore the strategies they employ to overcome these challenges. The research is guided by the following questions:

1. What challenges do students encounter in developing e-portfolios in public speaking courses?
2. What strategies do students use to address these challenges?

Recent studies (Bone et al., 2020; Scully et al., 2018) demonstrate the effectiveness of e-portfolios in enhancing students' engagement, self-reflection, and performance in language learning courses. The advantages of using e-portfolios in the teaching and learning process have been widely discussed in previous research. Yastibas and Cepik (2015) found that while teachers were supportive of e-portfolios in speaking classes, students reported difficulties with the process, particularly its complexity, which caused frustration for some. Barrot (2016) reported that e-portfolios enabled students to evaluate their own learning, monitor progress, and increase their motivation to learn.

Similarly, Siqueiros Quintana et al. (2022) and Ali et al. (2024) showed that an e-portfolio fostered positive washback by supporting professional development, encouraging critical thinking, and facilitating peer feedback. However, SARAÇ et al., (2022) cautioned that e-portfolios may also generate negative washback, such as anxiety about wider exposure of their work, ICT-related concerns, and challenges with independent learning. Lam (2022) further explained that students' success with e-portfolios is significantly influenced by their self-regulated learning practices. The use of e-Portfolio-based learning in speaking courses has also been addressed in earlier studies. For example, López-Crespo et al (2022) investigated the effectiveness of online learning designs incorporating e-portfolios to enhance self-efficacy in speaking skills. Their findings showed that e-portfolio-based learning had a significant impact on students' self-efficacy in speaking.

Despite these insights, few studies have examined the specific challenges students face and the strategies they employ when creating e-portfolios in speaking classes. To address this gap, the present study aims to explore the

difficulties and approaches experienced by students in developing e-portfolios for English language learning, particularly in speaking courses. By understanding these challenges and strategies, educators and learners can design more effective solutions to enhance the implementation of e-portfolios in speaking instruction.

Several classrooms at Universitas Ahmad Dahlan have adopted an e-portfolios as learning resources, particularly in the English education study program, including public speaking courses. Since the beginning of the 2022–2023 academic year, students in the classes have created and utilized e-portfolios through Google Sites. This platform is considered more productive and efficient for e-portfolio creation, as it is user-friendly and consumes less internet data (Olstad, 2020). Furthermore, Google Sites offers a range of features that allow users to embed text, graphics, photos, audio, video, and other materials pertaining to public speaking topics.

Nonetheless, e-portfolios are increasingly used effectively in public speaking courses, some students remain unfamiliar with the approach and naturally encounter challenges when developing them. This study, therefore, seeks to identify the challenges students face in public speaking classes when creating e-portfolios and to explore the strategies they employ to overcome these obstacles.

This study is framed by self-regulated learning theory (Zimmerman, 2000) and sociocultural theory (Vygotsky, 1978). Self-regulated learning theory explains how learners plan, monitor, and evaluate their own learning processes, making it relevant to understanding how students manage the difficulties e-portfolio development. Meanwhile, sociocultural theory emphasizes the importance of social interaction and mediation in learning, providing a lens to analyze how peer collaboration and lecturer support facilitate students' e-portfolio practices. Together, these theories provide a critical dimension of e-portfolio in the context of public speaking.

Gaps from Previous Studies

Previous research has documented several benefits of using e-portfolios, such as fostering learner autonomy, enhancing reflection, and promoting assessment transparency (Lam, 2022). At the same time, studies have identified disadvantages, including technical barriers, motivational challenges, and limitations in institutional infrastructure. However, few studies have specifically examined how Indonesian EFL students experience these challenges when creating e-portfolios in public speaking courses. Most prior research have focused on writing or general English skills, leaving a gap in understanding how students cope with technical barriers, time management difficulties, and creativity-related issues in speaking-focused courses. This study seeks to address this gap by investigating both the struggles and the strategies of Indonesian students in developing e-portfolios. In doing so, it contributes to practical pedagogy by offering insights for course design and to theoretical development by extending e-portfolio research into the context of public speaking.

METHODS

This study utilized a descriptive qualitative approach to examine the challenges students faced in developing e-portfolios and the strategies they employed to address these challenges in a public speaking class. Descriptive qualitative research focuses on analyzing and interpreting non-numeric data to provide an in-depth understanding of individuals' experiences and perspectives. This method was chosen to capture the complexities of the students' interactions with e-portfolios as a learning tool and to gain insights into their challenges and problem-solving strategies.

Setting and Participants

The research was conducted in Public Speaking class of the English Education Study Program at a university in Yogyakarta, Indonesia during the 2022/2023 academic year. The class consisted of 22 students, including 18 females and 4 males, all of whom were actively involved in e-portfolio development as part of their coursework. Eight participants (three males and five females) were purposively selected for interviews to ensure data saturation and provided diverse insights into their experiences with e-portfolio implementation. All participants had completed at least four prior English skills courses (Listening, Speaking, Reading, and Writing) before enrolling in Public Speaking, ensuring they had sufficient background in English learning.

Most students came from Indonesian-medium high schools, where English was taught as a foreign language, making their engagement with digital e-portfolio tasks particularly novel. Ethical approval for this study was obtained from the university's Institutional Review Board. Informed consent was sought via WhatsApp messages sent to all enrolled students, inviting them to participate in interviews. Only those who explicitly agreed were included.

Data Collection Techniques

Data were collected through three primary methods: observation, interviews, and documentation. Non-participant observation was conducted to record students' behaviors, challenges, and strategies during classroom activities without researcher interference. Semi-structured interviews were then carried out with students who had voluntarily provided consent via WhatsApp. The interview protocol allowed flexibility and follow-up questions, fostering interactive and dialogic exchanges in which participants could share their experiences openly while the researcher probed for clarification. In addition, documentation of students' e-portfolios and classroom activities was gathered to complement and validate the data from observations and interviews. All data sources were analyzed thematically, with coding informed by the theoretical frameworks of sociocultural theory and self-regulated learning. This ensured that both individual regulation and social interaction were captured in the analysis.

Research Instruments

The instruments used in this study included observation notes, an interview guideline, and documentation. Observation notes were utilized to systematically record students' behaviors and challenges during the development of their e-portfolios. The interview guide, designed with semi-structured questions,

ensured the collection of relevant and in-depth information aligned with the research objectives. Documentation consisted of students' e-portfolio work and other classroom artifacts, which served as supplementary evidence to triangulate the data. Sample interview questions included: "What challenges did you experience while developing your e-portfolio for the Public Speaking course?" and "How did you try to overcome those challenges?" Observation checklist items covered aspects such as students' participation in class activities, use of digital tools, and responses to peer feedback.

Trustworthiness

To ensure the credibility and trustworthiness, several strategies were implemented. Data triangulation was achieved by combining information from observations, interviews, and documentation. Peer review and expert feedback were sought to enhance the reliability of the findings. Thick description was used to provide a detailed account of the research setting and participants' experiences, allowing readers to understand the context and authenticity of the findings. Trustworthiness was also strengthened through social validation processes, including peer debriefing, consultations with colleagues, and member checking by sharing preliminary interpretations with participants for clarification.

Data Analysis Techniques

The data analysis process followed [Miles and Huberman's framework \(1984\)](#), which involves three main steps: data reduction, data display, and conclusion drawing. Data reduction included transcribing interviews, summarizing observation notes, and organizing documentation to identify relevant themes. Data display was conducted by systematically presenting themes in a structured format, enabling for comparison of students' challenges and strategies. Finally, conclusions were drawn by interpreting patterns in the data and verifying them through triangulation to ensure accuracy and depth. Thematic analysis was applied during coding to generate categories and themes from transcripts and observation notes.

Two researchers independently coded the data, and discrepancies were resolved through discussion until consensus was reached. Intercoder reliability was established through iterative discussions, and the final themes were validated by cross-checking with documentation and peer debriefing. In line with the frameworks, codes and themes were interpreted through the lenses of self-regulated learning ([Zimmerman, 2000](#)) and sociocultural theory ([Vygotsky, 1978](#)). This dual perspective enabled the analysis to capture both individual self-regulation and the social interactions shaping students' strategies, reflecting the social-process nature of qualitative inquiry as described in the data collection section.

RESULTS AND DISCUSSION

The study revealed a range of challenges faced by students in developing e-portfolios in the Public Speaking class. An initial and significant challenge was confusion about the concept of e-portfolios and how to develop them. Many

students were unfamiliar with this learning tool and lacked clarity about its purpose and format. This gap in understanding required a substantial period of adaptation as they navigated the new requirements. One student remarked, "At first, I didn't know what an e-portfolio was supposed to look like. I had no idea where to start, and it felt overwhelming." The complexity of developing e-portfolios created a learning curve, which was particularly daunting given the emphasis on independent learning. This contrasted sharply with the structured guidance students were accustomed to in traditional classroom practices.

The findings also highlighted technical and infrastructural barriers that hindered students' progress. A major obstacle was dependence on stable internet connectivity, which proved challenging for students in areas with unreliable network coverage. As one respondent stated, "Sometimes, I couldn't access the platform because the internet in my area was unreliable. It made it really hard to complete my work on time." In addition, limited technological proficiency posed another difficulty, as not all students were confident in using digital tools such as Google Sites, the primary platform for creating e-portfolios. These barriers frequently delayed progress and increased stress, making implementation more demanding than anticipated. Students also experienced compatibility issues when integrating multimedia elements such as videos and images into their e-portfolios, further complicating efforts to present a cohesive and dynamic digital project.

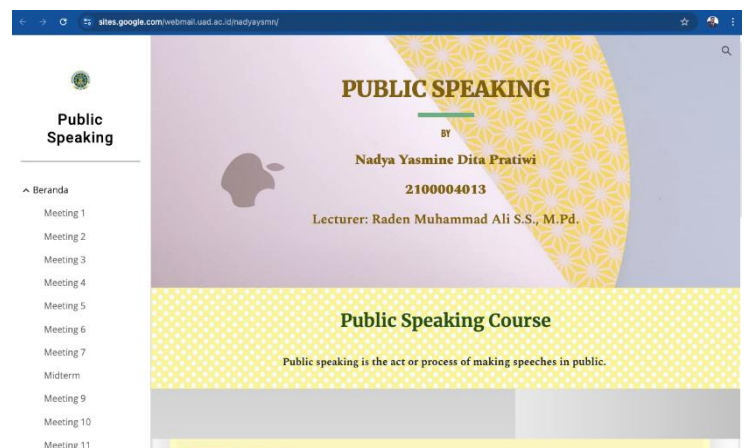


FIGURE 1 | Cover page of a public speaking student's e-portfolio, showing the initial layout and structure that illustrates how students organize the course identity and materials for each meeting

Time management emerged as a critical issue for students (see [Figure 2](#)). Developing of e-portfolios required meticulous planning and execution, ranging from collecting and organizing materials to reflecting on and showcasing their learning journey. Students reported that this process was highly time-intensive and often conflicted with their other academic responsibilities. As one participant noted, "I had to juggle between assignments from other courses and the time-consuming process of designing the e-portfolio,".

Procrastination and lack of motivation further exacerbated the problem, as many students struggled to allocate sufficient time and effort to create quality e-portfolios. This underscored the importance of fostering stronger time-management skills

to help students effectively meet the demands of e-portfolio development. Many students also noted that balancing multiple deadlines across different subjects alongside e-portfolio creation led to mental fatigue, which negatively affected their overall performance.

Another significant challenge was the demand for creativity and critical thinking in e-portfolio development. Students were required to present their learning outcomes in a visually appealing and coherent manner, requiring both aesthetic awareness and technical knowledge. One student confessed, "I'm not a creative person, so making the portfolio look good was very stressful for me." For many, this was an unfamiliar task that pushed them beyond their comfort zones. The requirement for originality and innovative thinking added additional pressure, as students had to ensure their e-portfolios stood out while still meeting the academic criteria set by their instructors. This often led to frustration, especially among those who lacked confidence in their creative abilities. Additionally, the overwhelming range of available for design and layout choices occasionally resulted in decision paralysis, adding another layer of difficulty to the process.

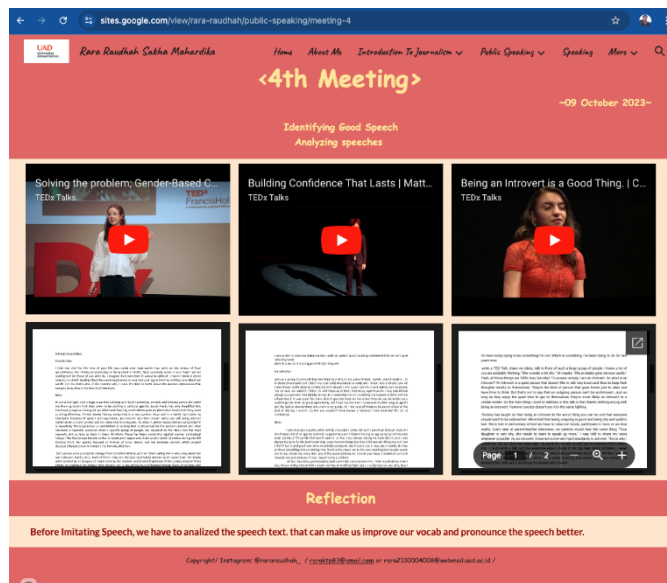


FIGURE 2 | A page of a student's e-portfolio containing learning materials, assignments, and reflection

The study also identified challenges related to balancing the academic and personal aspects of e-portfolio development. Students noted that creating an effective e-portfolio required more than simply fulfilling academic requirements; it demanded deep reflection on their learning journey, involving introspection and connections to personal experiences. As one student explained, "I struggled to find the right balance between making it professional and adding a personal touch." This blending of academic rigor with personal storytelling posed difficulties for students unaccustomed to reflective practices, as illustrated in [Figure 2](#), which shows a student's e-portfolio page combining assignments with reflective narratives. Integrating personal narratives often took longer than expected, compounding existing time-management challenges. Furthermore, some students struggle to identify meaningful moments or evidence to include in their e-portfolios, often feeling that their

contributions lacked significance or depth.

Despite these challenges, students adopted several strategies to overcome them. Collaborative learning emerged as a crucial approach, with many students forming study groups to share ideas, resources, and technical tips. One respondent shared, "Working with friends helped me understand how to organize my portfolio and gave me ideas I hadn't thought of before." Peer discussions not only demystified the e-portfolio development process but also fostered a sense of community, reducing feelings of isolation that some students initially experienced. Guidance from lecturers further provided valuable feedback, enabling students to refine their work and address specific difficulties effectively. Collaborative brainstorming sessions often generated innovative ideas, while peer encouragement motivated students to persist in their efforts.

Self-directed learning also played a pivotal role in navigating the complexities of e-portfolio development. Many students relied on online tutorials, such as YouTube videos and instructional websites, to acquire the necessary technical skills. One participant stated, "I watched a lot of tutorials online to learn how to use the platform and make my portfolio more interactive." This approach not only enhanced their digital proficiency but also fostered independence and resilience. Constructing concept maps was another effective strategy, helping students organize ideas and structure their e-portfolios logically. By visualizing their plans, students could align their work with expected outcomes and streamline the development process. Self-directed learning allowed them to explore new tools and features, adding distinctive elements that enriched their portfolios.

To address the issue of internet connectivity, students sought out locations with reliable access, such as campus facilities or public Wi-Fi hotspots. This proactive approach ensured smooth uploading of materials, collaboration with peers, and timely receipt of feedback. As one respondent explained, "I often stayed late on campus just to use the stable internet for uploading my videos." Furthermore, students recognized the importance of avoiding procrastination and implemented strategies to maintain discipline, such as setting personal deadlines, breaking tasks into manageable segments, and making steady progress. Daily or weekly schedules specifically dedicated to e-portfolio work proved particularly effective in reducing last-minute stress and ensuring consistent advancement.

Feedback and mutual support within the classroom environment emerged as a key enabler for success. Students frequently utilized tools such as Google Docs for collaborative editing and peer review, which facilitated constructive feedback and collective learning. As one participant explained "Having classmates review my work helped me see things I needed to fix that I hadn't noticed before." This interactive process not only enhanced the quality of the e-portfolios but also fostered a supportive learning culture. Regular feedback from lecturers helped students identify areas for improvement, encouraged the adoption of a growth mindset, and ultimately boosted their confidence and motivation. Structured in-class feedback sessions provided opportunities for students to share their progress and receive diverse perspectives, thereby enriching

their portfolios.

The psychological strategies also played a significant role in overcoming barriers. Many students emphasized building a sense of responsibility and accountability toward their e-portfolios. Viewing their portfolios as personal achievements rather than merely academic requirements motivated them to invest greater effort. One respondent noted, "I told myself this is something I can show off later, so I wanted it to be perfect." Celebrating small milestones and reflecting on their progress provided a sense of accomplishment, sustaining engagement over time. Moreover, students reported channeling frustrations into constructive actions, such as seeking inspiration from exemplary e-portfolios or using motivational tools like reward systems to maintain focus. Developing pride in their work proved to be a powerful driver of persistence and quality.

Institutional support was also emphasized as crucial in alleviating challenges. Workshops and training sessions conducted by lecturers equipped students' foundational skills for e-portfolio development, while also offering a platform to clarify doubts and receive hands-on guidance. One participant remarked, "The workshop really helped me understand what the lecturers were expecting from us." Access to digital infrastructure, such as computer labs and technical assistance, further supported students in managing the technological aspects of portfolio creation. Institutional interventions underscored the importance of collaborative efforts in fostering student success, with templates and sample e-portfolios serving as valuable references.

Cultural and social influences further shaped students' approaches to e-portfolio development. Peer pressure and competitive environments had both positive and negative impacts. While some students were motivated by their peers' progress and achievements, others experienced anxiety and self-doubt. One student reflected, "Seeing my classmates' portfolios sometimes made me feel like mine wasn't good enough, but I tried to focus on improving at my own pace." To counteract these effects, students were encouraged to prioritize personal growth and view the e-portfolio as a tool for self-expression rather than competition. This perspective helped them channel their energy into creating authentic and meaningful portfolios. Cultivating an environment of encouragement rather than comparison further reduced stress and fostered a healthier learning atmosphere.

The research also highlighted the long-term benefits of e-portfolio development. Beyond fulfilling academic requirements, students reported that the process enhanced their self-awareness and ability to articulate their skills and experiences. As one participant reflected, "Looking back at my portfolio, I can see how much I've grown, and it makes me proud." These attributes were deemed valuable for future endeavors, such as job applications and professional development. By documenting their learning journey, students created a repository of evidence that could be revisited and updated, positioning e-portfolios as dynamic tools for lifelong learning. This forward-looking perspective reinforced the relevance and utility of e-portfolios in modern education. Students also noted that reflecting on their accomplishments through e-portfolios boosted their confidence and strengthened their sense of achievement.

E-portfolios have become an increasingly popular tool in modern education, offering students a platform to showcase their learning journey and achievements. However, implementing e-portfolios into the academic curriculum presents both challenges and opportunities. This discussion examines the findings related to the challenges students encounter in developing e-portfolios and the strategies they employ to overcome them. By addressing these aspects, educators and institutions can gain valuable insights to optimize e-portfolio integration and ensure it becomes a transformative learning experience. The discussion also situates the findings within existing literature, highlighting practical solutions and implications for future practice. Notably, these findings align with prior studies in Indonesia ([Muin et al., 2021](#)) and Asia ([Alshahrani et al., 2023](#)), which similarly reported difficulties with technical and motivational aspects of e-portfolio development. Such consistency suggests that the challenges faced by Indonesian students are not isolated but reflect broader regional trends in implementing digital portfolio practices.

Challenges and Barriers in E-Portfolio Development

The findings of this study revealed multiple challenges students faced in the development of e-portfolios, ranging from technical obstacles to the demand for creativity. Confusion regarding the very concept of e-portfolios highlighted the necessity for preliminary orientation sessions. Students' remarks, such as "I didn't know where to start," indicate a knowledge gap that educators must address. Consistent with recent findings by [Ismail \(2023\)](#), the initial complexity of e-portfolios can lead to frustration unless students are provided with clear guidelines and examples. Effective instructional scaffolding, coupled with well-designed workshops, can significantly mitigate this challenge by offering step-by-step support.

Technical difficulties, such as unreliable internet connectivity and insufficient technological skills, also impeded progress. One respondent noted, "The internet in my area made it difficult to finish tasks on time." Similar challenges were documented by [Alshahrani et al. \(2023\)](#), who emphasized the importance of infrastructure support in successful e-portfolio implementation. Institutions should therefore prioritize providing reliable digital infrastructure, such as campus-wide high-speed internet, and alongside training sessions tailored to varying levels of digital literacy. Additionally, platforms selected for e-portfolios development should be user-friendly to minimize the learning curve.

Time management emerged as another recurring challenge, highlighted by a participant's comment: "Balancing e-portfolios and other assignments was exhausting." This finding aligns with [Khound et al. \(2024\)](#), who suggested that time-intensive activities like e-portfolio development demand strong organizational skills. Educators can support students by integrating project management tools into the curriculum, fostering effective planning practices, and setting incremental deadlines. Breaking larger tasks into smaller, manageable segments can help students maintain steady progress and reduce last-minute pressures.

Strategies for Overcoming Challenges

Despite these challenges, students employed various

strategies to mitigate difficulties in e-portfolio development. Collaborative learning emerged as a pivotal strategy. Respondents highlighted the value of peer discussions, with one noting, "Working with friends helped me understand how to organize my portfolio." Collaborative learning aligns with Vygotsky's theory of social constructivism, which posits that peer interactions facilitate cognitive development. Incorporating collaborative activities, such as group reviews and peer evaluations, into e-portfolio assignments can therefore enhance learning outcomes. Platforms like Google Docs further support real-time collaboration, enabling students to share feedback and ideas seamlessly. Furthermore, peer mentoring programs that pair students with varying levels of expertise could promote mutual learning and skill development.

Self-directed learning was another prominent strategy. One participant explained: "I watched tutorials to make my portfolio interactive.", highlighting reliance on online resources. This approach reflects Knowles' concept of andragogy, which emphasizes learner autonomy and initiative (Cleavenger, 2020). Similarly, Othman et al. (2018) found that self-directed learning not only enhances technical skills but also strengthens resilience and adaptability. To encourage this, educators could curate lists of reliable online resources and design assignments that promote independent exploration, such as tasks requiring students to "Find and integrate a new feature into your portfolio.". Providing access to technical support and online discussion forums would further empower students to troubleshoot challenges and develop autonomy in their learning.

Enhancing Creativity and Reflective Practices

Developing a visually appealing and coherent e-portfolio posed significant challenges, particularly for students who doubted their creative abilities. One participant shared "I'm not a creative person,", illustrating a psychological barrier. Creativity is a crucial component of e-portfolio development, often pushing students beyond their comfort zones. While their presents challenges, it also offers opportunities for growth. Rowley and Munday (2022) highlight that e-portfolios are inherently creative tasks, requiring a balance between aesthetic appeal and content relevance.

To address this challenge, educators can incorporate design-thinking workshops into the curriculum, offering hands-on opportunities to for students to develop skills in layout design, multimedia integration, and storytelling. Such workshops can nurture creativity and critical thinking, helping students build confidence in their ability to communicate ideas effectively through digital formats. Presenting exemplary e-portfolios can also inspire students and set benchmarks for quality. Exposure to diverse models helps students visualize success and recognize the various ways multimedia elements can enhance their work.

Reflective practices during the design process play an equally vital role. Structured reflection enables students to assess their progress, refine approaches, and align their portfolios with both academic and personal goals. Encouraging constructive peer feedback further strengthens reflective learning by providing new perspectives, identifying weaknesses, and refining creative outputs. As Hui et al. (2023) suggest, peer input fosters collaboration and collective

improvement, ultimately leading to more polished and impactful e-portfolios.

The reflective practices reported by students also connect to theories of self-regulated learning and metacognition. By setting personal goals, monitoring progress, and adjusting strategies, students demonstrated elements of self-regulation—skills identified by Lam (2022) as critical for successful independent learning. These processes indicate that e-portfolio development not only enhances technical competence but also cultivates higher-order metacognitive abilities, reinforcing its value as a transformative pedagogical tool.

Institutional Support and Long-Term Benefits

Institutional support played a crucial role in alleviating challenges. Workshops, as highlighted by a respondent's remark "The workshop clarified what was expected" proved instrumental in equipping students with essential skills. As Torre (2019) argues, structured training sessions are vital for ensuring that students grasp both the technical and conceptual aspects of e-portfolio development.

Beyond workshops, institutions must also invest in digital infrastructure. Reliable internet access, well-equipped computer labs, and readily available technical assistance are essential for successful implementation. Templates and sample portfolios provided by instructors serve as practical references, bridging the gap between abstract expectations and tangible outcomes. Such initiatives not only enhance students' experiences but also highlight the institution's commitment to fostering digital literacy and long-term professional growth.

7. Rara Raudhah Sakha Mahardika (2100004008)

URL e-Portfolio: <https://sites.google.com/view/rara-raudhah/home>

Members of The Groups :

5. M. Ghoni Raihan

6. Azzam Firdausi

7. Rara Raudhah Sakha Mahardika

Meeting	Topic	Comment, question, advice, etc (Write your first name: your comment) Example: Raden : The topic of your speech is awesome. Can you tell me how to find such an interesting topic?	Response
1	Introduction about Public Speaking Course	Raihan: I got a lot of information about public speaking after visiting your page, Ra. The display also looks fascinating! Great job! Azzam: You put the information clearly there, Rara. You attached the Lesson plan as well. Since it is the introduction session, I think it is pretty comprehensive information and a clear strong bridge to Public Speaking. The view and the display are eye-catching and not boring, good job!	Rara: Glad to hear that, Han. thank you! Rara: Thank you, Azzam, I appreciate it so much.
2	What is Public Speaking and How We Do it?	Raihan: Actually, I'm always nervous to speak in front of others. And after visiting your page, I learned how to do it little by little! Thank you so much, Rara! Azzam: This one as well. You wrote comprehensively great information there. You put the explanation and description as well. I hope everyone who visits your page get the information they want. Well done, Rara! And I hope i could do the public speaking confidently!	Rara: Hope you can improve your speaking skills! Thank you, Han! Rara: It is my pleasure knowing that, Azzam. Thank you so much. I, pretty sure you can do it one day, Zam!
3	Preparing Imitating Speech	Raihan: You chose good topics and speech there ra ! love it Azzam: You served a quite clear information there, Ra! I never knew that imitating or impersonating could help my public speaking. Thank you!	Rara: Thanks zam for your comment!

FIGURE 3 | A screenshot of the students' peer discussion forum, illustrating how collaborative learning and peer feedback supported the development of their e-portfolios

These findings carry implications for curriculum design. Lecturers should scaffold e-portfolio tasks gradually, beginning with structured guidelines and examples before progressing to more open-ended assignments that encourage creativity. Systematic integration of peer collaboration, for instance, through peer review sessions or collaborative workshops, can also help students overcome motivational and technical barriers while enhancing reflective practices.

Beyond their academic utility, e-portfolios were also perceived as valuable tools for professional development. As one participant reflected, "Looking back at my portfolio, I can see how much I've grown." This observation resonates with [Mapundu and Musara \(2019\)](#), who described e-portfolios as lifelong learning tools that document growth and achievements. Encouraging students to view their e-portfolios as evolving documents prepares them for future applications, such as job interviews or graduate studies. Incorporating career-oriented elements, such as a Skills and Achievements section, further enhances their relevance. By framing e-portfolios as assets for professional advancement, educators can inspire students to invest greater effort and creativity in their development.

This study is limited by its small sample size and focus on a single institution, which restricts the generalizability of the findings. Future research could adopt longitudinal designs to examine how students' reflective practices and strategies evolve over time. Comparative studies across institutions and skills areas (e.g., writing, listening, or interdisciplinary projects) would also provide deeper insights into the transferability of e-portfolio practices.

CONCLUSION

The findings of this study emphasize that while the development of e-portfolios in public speaking classes is highly beneficial, it also presents a range of interconnected challenges for students. Technical difficulties, such as unstable internet connections and limited technological proficiency, were compounded by psychological barriers, including low motivation and creative struggles. These challenges underscore the need for structured guidance and sustained support. At the same time, students' adaptive strategies, such as collaborative learning, self-directed exploration, and effective time management, reflect their resilience in navigating obstacles. This highlights the importance of fostering a supportive learning environment where students have access to resources, constructive feedback, and opportunities for skill enhancement.

To ensure the successful integration of e-portfolios, institutions must prioritize both technical and pedagogical support. Workshops, reliable infrastructure, and clear guidelines are crucial in reducing students' anxiety and confusion. Educators can further encourage reflective practices and provide templates that stimulate creativity. By systematically addressing challenges and promoting both collaborative and independent learning strategies, e-portfolios can serve as a transformative educational tool. Beyond enhancing academic outcomes, they also cultivate essential skills such as critical thinking, time management,

and self-motivation, competencies that are invaluable for students' future careers and lifelong learning.

Nevertheless, this study has certain limitations, including a relatively small sample size (22 students from a single university) and its focus on a single English skill (public speaking), which may restrict the generalizability of the findings. Despite these limitations, the study offers valuable contributions: practically, it informs teachers and institutions about effective ways to integrate e-portfolios; methodologically, it demonstrates the application of qualitative approaches to exploring student experiences; and theoretically, it enriches understanding by drawing on sociocultural theory and self-regulated learning theory within the context of e-portfolio development. For future research, studies involving larger and more diverse samples, comparative investigations across different English skills, and the use of mixed-methods designs are recommended to strengthen the evidence base and broaden insights.

ACKNOWLEDGEMENTS

This program is financed by the Directorate of Research, Culture, and Community Service, Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia, under the 2024 fiscal year and The Bureau of Research and Community Service, Universitas Ahmad Dahlan, fiscal year 2024. The authors also acknowledge the use of ChatGPT (OpenAI) to assist in improving the clarity and fluency of the English language in this manuscript. The authors take full responsibility for the final content of the article.

REFERENCES

- Ali, R. M., Hastuti, D., Ananda, A. R., & Biddinika, M. K. (2024). Learning speaking using Google Sites-based e-portfolio: Indonesian students' experience. *Exposure Journal*, 197(1), 197–214. <https://journal.unismuh.ac.id/index.php/exposure>
- Alshahrani, S. M., Mohamed, H., Mukhtar, M., & Mokhtar, U. A. (2023). The adoption of the e-portfolio management system in the Technical and Vocational Training Corporation (TVTC) in Saudi Arabia. *International Journal of Information Management Data Insights*, 3(1), 100148. <https://doi.org/10.1016/j.ijime.2022.100148>
- Aulia, D., Yulastri, D., & Handayani, W. (2018). The use of e-portfolio for improving students' English skills: A pilot study in English Department, the State Polytechnic of Padang. *Journal Polingua: Scientific Journal of Linguistic Literatura and Education*, 5(2), 52–59. <https://doi.org/10.30630/polingua.v5i2.37>
- Barrot, J. S. (2016). Using Facebook-based e-portfolio in ESL writing classrooms: Impact and challenges. *Language, Culture and Curriculum*, 29(3), 286–301. <https://doi.org/10.1080/07908318.2016.1143481>
- Bolliger, D. U., & Shepherd, C. E. (2010). Student perceptions of e-portfolio integration in online courses. *Distance Education*, 31(3), 295–314.

- <https://doi.org/10.1080/01587919.2010.513955>
- Bone, E., Greenfield, R., Williams, G., & Russell, B. (2020). Proceedings of the ASCILITE 2020 Conference (pp. 194–200). ASCILITE.
- Ciesielkiewicz, M. (2019). Education for employability: The e-portfolio from school principals' perspective. *On the Horizon*, 27(1), 46–56. <https://doi.org/10.1108/OTH-01-2019-0001>
- Cleavenger, J. M. (2020). Missing and murdered Indian women in Montana. [Unpublished report].
- Hui, L., Ippolito, K., Sarsfield, M., & Charalambous, M. (2023). Using a self-reflective e-portfolio and feedback dialogue to understand and address problematic feedback expectations. *Assessment and Evaluation in Higher Education*, 48(1), 1–14. <https://doi.org/10.1080/02602938.2023.2232960>
- Isiyaku, D. D., Ayub, M. A. F., & AbdulKadir, S. (2018). Antecedents to teachers' perceptions of the usefulness of ICTs for business education classroom instructions in Nigerian tertiary institutions. *Asia Pacific Education Review*, 19(3), 337–352. <https://doi.org/10.1007/s12564-018-9525-x>
- Ismail, S. S. (2023). Barriers towards the implementation of e-portfolio in education based on the diffusion of innovation theory. *International Journal of Learning, Teaching and Educational Research*, 22(4), 512–540. <https://doi.org/10.26803/ijlter.22.4.28>
- Khound, M., Das, B. K., & Shankar, K. J. (2024). Development and validation of e-portfolio for undergraduate medical students in pediatrics. *Indian Pediatrics*, 61(8), 740–746.
- Lam, R. (2022). E-portfolios for self-regulated and co-regulated learning: A review. *Frontiers in Psychology*, 13, 1–8. <https://doi.org/10.3389/fpsyg.2022.1079385>
- Lewis, L. (2017). E-portfolio as pedagogy: Threshold concepts for curriculum design. *E-Learning and Digital Media*, 14(1–2), 72–85. <https://doi.org/10.1177/2042753017694497>
- López-Crespo, G., Blanco-Gandía, M. C., Valdivia-Salas, S., Fidalgo, C., & Sánchez-Pérez, N. (2022). The educational e-portfolio: Preliminary evidence of its relationship with students' self-efficacy and engagement. *Education and Information Technologies*, 27(4), 5233–5248. <https://doi.org/10.1007/s10639-021-10827-2>
- Mapundu, M., & Musara, M. (2019). E-portfolios as a tool to enhance student learning experience and entrepreneurial skills. *South African Journal of Higher Education*, 33(6), 191–214. <https://doi.org/10.20853/33-6-2990>
- Miles, M. B., & Huberman, A. M. (1984). Drawing Valid Meaning from Qualitative Data: Toward a Shared Craft. *Educational Researcher*, 13(5), 20–30. <https://doi.org/10.3102/0013189X013005020>
- Muin, C. F., Hafidah, H., & Daraini, A. M. (2021). Students' perceptions on the use of e-portfolio for learning assessment. *Al-Ishlah: Jurnal Pendidikan*, 13(1), 497–503. <https://doi.org/10.35445/alishlah.v13i1.485>
- Nagle, L., O'Connell, M., & Farrelly, T. (2019). A gap in governance: Acknowledging the challenges of organic e-portfolio implementation. *Educational Media International*, 56(4), 328–342. <https://doi.org/10.1080/09523987.2019.1682271>
- Olstad, H. A. (2020). E-portfolio introduction: Designing a support process. In *Lecture Notes in Computer Science* (Vol. 12518, pp. xxx–xxx). Springer International Publishing. https://doi.org/10.1007/978-3-030-63212-0_17
- Othman, R., Kamal, N. M., Alias, N. E., Ismail, S., & Sahiq, A. N. M. (2018). Positive psychological traits and career adaptability among millennials. *International Journal of Academic Research in Business and Social Sciences*, 8(9), 1–12. <https://doi.org/10.6007/ijarbss/v8-i9/4706>
- Rahman, A., & Sari, D. P. (2023). Digital e-portfolio implementation in Indonesian higher education: Challenges and opportunities. *Indonesian Journal of Applied Linguistics*, 13(2), 301–315. <https://doi.org/10.17509/ijjal.v13i2.XXXX>
- Rowley, J., & Munday, J. (2022). An arts-based learning model: Synergies of artist mentorship, e-portfolio and discovering 'possible self.' *The Australian Journal of Indigenous Education*, 51(1), 1–16. <https://doi.org/10.55146/ajie.2022.30>
- Saraç, M., Saraç, N., & Öztürk, M. (2022). A literature review on e-portfolios in teaching writing in English as a foreign language classes. *RumeliDE Dil ve Edebiyat Araştırmaları Dergisi*, 9(29), 904–916. <https://doi.org/10.29000/rumelide.1136600>
- Scully, D., O'Leary, M., & Brown, M. (2018). The learning portfolio in higher education: A game of snakes and ladders. In *Proceedings of the Learning Portfolio in Higher Education Conference* (March). <https://doi.org/10.13140/RG.2.2.18883.71208>
- Siqueiros Quintana, M. G., Mungarro Robles, G. del C., & Robles Tarazón, H. S. (2022). The e-portfolio as a learning evaluation tool: Main benefits and challenges. In S. Hosseini, D. H. Peluffo, J. Nganji, & A. Arrona-Palacios (Eds.), *Technology-enabled innovations in education* (pp. 199–207). Springer Nature Singapore. https://doi.org/10.1007/978-981-16-XXXX-X_17
- Strohmeier, S. (2010). Electronic portfolios in recruiting? A conceptual analysis of usage. *Journal of Electronic Commerce Research*, 11(4), 268–278.
- Syzdykova, Z., Koblandin, K., Mikhaylova, N., & Akinina, O. (2021). Assessment of e-portfolio in higher education. *International Journal of Emerging Technologies in Learning*, 16(2), 120–134. <https://doi.org/10.3991/ijet.v16i02.18819>
- Tan, J., & Lee, M. H. (2024). Exploring metacognitive engagement through e-portfolios in Asian EFL contexts. *Asia-Pacific Education Researcher*, 33(1), 77–92. <https://doi.org/10.1007/s40299-023-00789-2>
- Torre, E. M. (2019). Training university teachers on the use of the e-portfolio in teaching and assessment. *International Journal of ePortfolio*, 9(2), 97–110. <http://www.theijep.com>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

- Yastibas, A. E., & Cepik, S. (2015). Teachers' attitudes toward the use of e-portfolios in speaking classes in English language teaching and learning. *Procedia – Social and Behavioral Sciences*, 176, 514–525. <https://doi.org/10.1016/j.sbspro.2015.01.505>
- Zheng, Y. (2022). Social media as an e-portfolio platform: Effects on L2 learners' speaking performance. *Language Learning & Technology*, 26(1), 1–19. <https://doi.org/10.10125/73400>
- Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). San Diego, CA: Academic Press. <https://doi.org/10.1016/B978-012109890-2/50031->

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright ©2025 Raden Muhammad Ali, Pamastu Narpaduita, Muhammad Iqwan Sanjani, Dwi Hastuti, Ummi Khaerati Syam, Muhammad Kunta Biddinika. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The influence of AI-Generated feedback on university students' debate skills

Landry Dwiyoga Daniswara*

Universitas Negeri Surabaya, Indonesia

The use of Artificial Intelligence (AI) in higher education is increasing, especially through AI-assisted feedback tools such as ChatGPT. These tools provide feedback not only on language use but also on reasoning and argument quality. This study examined the effect of ChatGPT-assisted feedback on students' debate performance using a one-group pretest-and-post-test pre-experimental design. Eight undergraduate students participated in simulated British Parliamentary debates, each debating two motions on the same topic. Their performance was assessed using an expert-validated debate rubric. Quantitative data were analyzed with SPSS version 29. The results showed that the average pretest score was 75.75 (SD = 1.282) and the post-test score was also 75.75 (SD = 1.282). A paired-samples t-test confirmed no significant improvement, $t(7) = 0.00$, $p = 1.00$, Cohen's $d = 0.00$, 95% CI [-0.63, 0.63]. However, a strong correlation between the two tests was found, $r = 0.826$, $p = 0.011$, showing that consistent student performance. Semi-structured interviews further revealed that many students relied passively on ChatGPT, often copying its suggestions directly instead of adapting them, with limited the possible benefits of the tool. These findings suggest that AI-assisted feedback alone does not guarantee performance gains; its effectiveness depends on students' critical engagement and integration of feedback into their debating practice. The study highlights the importance of teacher guidance in supporting effective use of AI feedback and suggests future studies with larger samples and extended training periods.

Keywords: Artificial Intelligence, argument, ChatGPT, debate, feedback

INTRODUCTION

Debate has become a key part of education because it helps students develop critical thinking skills ([Baziganya, 2024](#); [Chen et al., 2022](#); [Chew et al., 2021](#); [Ko & Mezuk, 2021](#); [Pahmi & Njah, 2021](#)). As a result, improving students' debate performance has been increasingly significant. Debating activities are now integrated into many subjects, especially in English language learning, where they serve as an effective teaching method. In universities, student debate clubs and specialized speaking courses further support this goal. At Universitas Negeri Surabaya (well-known as UNESA), for example, a debate community provides students from various academic backgrounds with opportunities to strengthen their critical thinking and communication skills. The community regularly organizes sessions, conduct expert workshops on argument construction, and encourages participation in debate competitions. However, despite these activities, limited research has examined how AI feedback can improve debate performance, especially for EFL students. Most studies focus on debate as a method to enhance language learning but do not investigate the specific impact of AI feedback on debating

OPEN ACCESS

ISSN 2503 3492 (online)

*Correspondence:

Landry Dwiyoga Daniswara
landrydaniswara@gmail.com

Received: 26th May 2025

Accepted: 8th September 2025

Published: 24th October 2025

Citation:

Daniswara, L.D. (2025). The influence of AI-Generated feedback on university students' debate skills. *JEES (Journal of English Educators Society)*, 10(2).
<https://doi.org/10.21070/jees.v10i2.1941>

This seeks to address this gap by examining students' debate performance in such an environment.

The problem of this study arises when students face several challenges in debating, particularly in producing language in English. Many Englishers as a Foreign Language (EFL) learners struggle with vocabulary and fluency, which hinder their ability to express arguments clearly and persuasively (Alasmari & Ahmed, 2012; Astuti et al., 2022). These language barriers often reduce their confidence, making them feel insufficient for meaningful participation (Ban et al., 2023; Sanonguthai, 2011). Another challenge is understanding the debate motion themselves. Debate topics can be complex and students with limited background knowledge or unfamiliarity with specialized vocabulary may find it hard to fully engage (Ban et al., 2023; Nurakhir et al., 2020). This lack of preparation can cause frustration and anxiety, which negatively impacts debate performance (Hernawan & Yusuf, 2021; Wariyati, 2023).

Critical thinking also plays a crucial role in debate quality. Students are expected to quickly develop well-reasoned arguments and counterarguments which can be mentally demanding (Sawyer, 2015; Ali et al., 2020). This pressure intensifies when they must engage critically with opposing views that they initially find intimidating (Lestari, 2018; Sen et al., 2023). According to Daniswara and Cahyono (2023), strong arguments should include Assertion, Reasoning, Evidence, and Link Back. When students fail to meet these criteria, it reveals gaps in their critical thinking skills.

To improve students' debate skills, educators use various teaching strategies. Spaska et al. (2021) note that methods such as modeling, guided practice, and collaborative learning can significantly enhance debate abilities. Regular practice builds students' confidence, sharpens their argumentation skills, and deepens their understanding of debate topics or motions. Recently, AI-powered tools have emerged as valuable resources in debate education. These tools provide access to vast information, curate relevant content, suggest reading materials, and even stimulate debate scenarios for practice in controlled environments (Tjahyaningtjas, 2023; Munir et al., 2022). Such application not only boosts critical thinking and analytical skills but also improve students' comprehension of debate topics. Additionally, AI can evaluate performances by identifying strengths and weaknesses, thereby helping students prepare more effectively (Ivanova & Petrova, 2023; Yang et al., 2021).

AI tools extend beyond supplying information. They generate diverse perspectives and counterarguments, encouraging students to critically evaluate their own positions. Research highlights the value of engaging with multiple viewpoints to build well-rounded arguments (Rahmi & Jamaluddin, 2021; Aisyah, 2024). Integrating AI systematically into debate exercises fosters deeper evidence analysis, clearer expression of ideas, and constructive discussions, skills that are essential both academically and in real-world contexts (Sugianti, 2024; Idroes, 2023). While existing studies have explored AI feedback in writing and language learning, few have examined how AI-generated feedback directly improve debating skills. This gap underscores the need for further research, and the present study seeks to contribute to this underexplored area.

Although earlier versions of ChatGPT faced challenges with accuracy and content quality (Liu et al., 2024), most recent versions, such as ChatGPT-4 and later, show significant improvements, providing more precise and reliable responses (Didonna et al., 2024; Pang et al., 2024). Research has shown ChatGPT's usefulness in English learning, providing study materials, acting as a conversational partner, and delivering instant feedback (Nugroho et al., 2023). Moreover, some studies suggest that ChatGPT can support improvement in overall language proficiency (Akopiants, 2023; Barabad & Anwar, 2024).

Due to its versatile capabilities, ChatGPT has been increasingly adopted in debate education (Wang et al., 2023). The introduction of ChatGPT-4 offers a valuable opportunity to enhance students' argumentative skills. As a generative AI model, ChatGPT-4 can assist learners at various stages of debate preparation, from idea generation and argument organization to providing constructive feedback on written work. This support is especially beneficial for EFL students, who often struggle to clearly express and structure their arguments in a second language. Recent studies indicate that ChatGPT positively impacts students' writing development, particularly in argumentative essay writing, which is a critical aspect of debating (Esmacil et al., 2023; Nguyen et al., 2024; Wang, 2024). The theoretical foundation for using AI feedback in this research draws on formative assessment, where learning improves through continuous and supportive feedback (Prompiengchai, Narreddy, & Joordens, 2025), and sociocultural theory, which sees learning as something socially mediated through interaction and scaffolding (Li, 2025). In this study, ChatGPT is viewed as a feedback partner that provides comments for reflection, helps students refine their reasoning, and guides them in organizing their arguments more clearly.

Using ChatGPT in English learning and debating can lead to two possible outcomes. On one hand, ChatGPT offers diverse perspectives that encourage critical thinking (Jiang & Hyland, 2024; Halaweh, 2023; Oon et al., 2023). In this case, students are inspired by ChatGPT input, which improve their debating skills. On the other hand, there is a risk of overreliance, where students depend too heavily on ChatGPT and limit their own critical engagement (Esmacil et al., 2023; Ruff et al., 2024). Therefore, thoughtful use of ChatGPT with proper guidance and moderation is essential to maximize its benefits while minimizing potential downsides.

Integrating AI in education serves several purposes. It can enhance classroom debates by providing real-time feedback and personalized learning experiences tailored to student's needs. AI can assess arguments, identify strengths and weaknesses, and offer constructive critiques to strengthen reasoning and presentation skills. Studies show that AI use in education promotes academic performance and engagement by supporting individualized learning (Onesi-Ozigagun, 2024; Trisnawati, 2023). Furthermore, AI helps students develop self-awareness of their strengths and areas for growth, fostering a growth mindset and resilience (Abbas et al., 2023; Xu, 2023).

Although many studies have examined AI's role in supporting debate, such as supplying information, helping with argument development, and providing feedback, there

remains limited evidence on how AI-generated feedback directly influences students' actual debate performance, especially among EFL university students. This study addresses this gap by testing the effect of ChatGPT feedback on students' debating skills in a real university setting.

METHODS

This study employed a mixed-methods exploratory design, combining a one group pretest-post-test pre-experimental design with interviews to strengthen and validate the findings through qualitative triangulation (Creswell & Creswell, 2018). A pre-experimental design was chosen because no control group was available, allowing for a comparison of participants' performance before and after the intervention. To gain a deeper understanding of the students' experiences and perspectives, semi-structured interviews were also conducted.

The participants were eight undergraduate students from UNESA, all active members of the university's debating union and representing diverse academic majors. Purposive sampling was used, as these students already had debating experience and familiarity with the British Parliamentary format. This ensured that they could engage meaningfully with the intervention and provide reliable reflections on their use of ChatGPT. The small sample size reflects the exploratory nature of the research, which aimed not to generalize findings but to capture insights into how trained debaters interact with AI-assisted feedback during debate preparation.

Data Collection

Quantitative Data Collection Process

The quantitative data consisted of pretest and post-test scores from the debate simulations. In both tests, eight students were assigned different motions to debate. The use of different motions for the pretest and post-test was justified on several grounds. First, both motions were selected from the same theme, education, to ensure thematic consistency. Second, using different motions helped prevent memorization bias, as repeating the same motion could lead participants to rely on prior preparation rather than demonstrating skill development. Third, the focus of the study was not solely on debate content but on complexity of argument traits, including argument structure, logical reasoning, elaboration, and responsiveness.

The motions used for the pretest and post-test simulations are presented below.

TABLE 1 | Debate motion for pretest and post-test

	Motion
Pretest	This House Would glorify the success of online learning
Post-test	This House Would bring back national exam

During pretest, students were not allowed to use ChatGPT and had to prepare their cases independently. The motion was either provided before the simulation or delivered impromptu. Following the British Parliamentary format, students were given 15 minutes for case-building. After the pretest scores were recorded, the intervention phase began. Participants

then attended three training sessions on how to use ChatGPT effectively for British Parliamentary debate preparation. Each session focused on different debating skills. The first session trained students to interpret motion types, parameters, and core clashes. The second meeting guided them in constructing arguments, developing extensions, and identifying relevant examples. The third session emphasized elaboration and rebuttals. Across these sessions, students practiced with guided prompts, reflected on useful outputs, and learned to critically evaluate ChatGPT's feedback. After the training, students were given one week to independently apply the same strategies when preparing for the post-test motion. This design aimed to discourage passive copying and ensure that ChatGPT was used as a feedback tool to strengthen argument quality.

In this study, feedback was provided through ChatGPT (GPT-4). Rather than focusing on grammar correction or language accuracy, the tool mainly supported students in the process of building arguments. The feedback addressed three main areas. First, it supported debate strategy, for example, by helping students interpret the motion and decide their team's position. Second, it assisted in argument development, such as creating points, adding extensions, and suggesting relevant examples. Third, it contributed to the integration of information, allowing ChatGPT's ideas to be adapted and applied in different ways. In this sense, ChatGPT functioned as a feedback partner, providing material that students could use to prepare their strategy and engage more actively in debates.

ChatGPT was selected for this study because of its practical benefits, particularly its capacity to enable active interaction through real-time question-and-answer exchanges. This immediate responsiveness was expected to enhance students' understanding by encouraging deeper engagement with dynamically accessible content (Gupta et al., 2024; Ifraheem et al., 2024).

TABLE 2 | Students' score of debate simulation

Student	Team	Pretest	Post-test
Prime Minister (PM)	Opening Government	76	76
Deputy Prime Minister (DPM)		74	75
Leader of Opposition (LO)	Opening Opposition	77	76
Deputy Leader of Opposition (DLO)		76	77
Member of Government (MG)	Closing Government	77	77
Government Whip (GW)		77	77
Member of Opposition (MO)	Closing Opposition	75	74
Opposition Whip (OW)		74	74

For the post-test, students were permitted to use ChatGPT to assist in constructing their arguments within the same 15 minutes preparation time. After the post-test, individual scores were assigned using marking scheme adapted from the Indonesia National University Debate Championship

(NUDC) guidebook, with scores ranging from 50 to 100 (see Appendix). The rubric was validated by two debate experts who confirmed its relevance for assessing argumentation, reasoning, evidence, and style. Both experts also had extensive experience adjudicating debate competitions.

The following section presents the results of the pretest and post-test assessments of the eight students who participated in the British Parliamentary debate simulation.

Qualitative Data Collection Process

After the post-test simulation, semi-structured interviews were conducted individually with each student to capture their experiences and reflections on using ChatGPT. The interview focused primarily on the kinds of information students sought from ChatGPT and how they applied it in debate preparation. Only two key questions were used, as the purpose was not to collect a wide range of responses but to provide triangulation for the quantitative data. Keeping the questions focused allowed students to give more detailed and meaningful reflections without unnecessary repetition. The questions were: (1) What kinds of information did you look for when using ChatGPT? (2) How did you use and integrate the information provided by ChatGPT? These two questions were deliberately chosen to capture both the input (the information students sought) and the processing (how they integrated it into their debating strategy). The insights gained from these interviews proved valuable for interpreting the pretest and post-test results, particularly in assessing whether students' performance showed measurable improvement.

Data Analysis

Quantitative Data Analysis

The pretest and post-test data were analyzed using SPSS version 29 to examine whether the use of ChatGPT in the post-test led to a statistically significant improvement in students' performance compared to their pretest results. Three primary outputs were utilized for interpretation: paired sample statistics, paired sample correlations, and paired sample t-tests. These outputs provided insights into the descriptive statistics, the mean differences between the two tests, the correlation between them, and the statistical significance of any observed differences.

Qualitative Data Analysis

The semi-structure interview data were analyzed using content analysis to uncover key insights and themes regarding students' use of ChatGPT. The process followed three main stages: inductive coding, code analysis, and interpretation. Specifically, the steps included: (1) collecting the interview data, (2) transcribing the recordings, (3) carefully reviewing the transcripts, (4) labeling relevant sections with codes and organizing them, and (5) interpreting the codes to identify descriptive themes, which were then presented in the study's findings.

RESULTS AND DISCUSSION

Test of Normality

Before conducting the paired-samples test, the normality of the data was assessed.

TABLE 3 | Test of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Pretest	.210	8	.200*	.843	8	.082
Post-test	.210	8	.200*	.843	8	.082

To test the assumption of normality, the Kolmogorov-Smirnov test with Lilliefors correction and the Shapiro-Wilk test was applied. The results showed that the pretest scores (Shapiro-Wilk, $p = 0.082$) and the post-test scores (Shapiro-Wilk, $p = 0.082$) followed a normal distribution ($p > 0.05$). The Kolmogorov-Smirnov test also supported this result ($p = 0.200$). Given the small sample size ($n = 8$), the Shapiro-Wilk test was used as the main reference. As the normality assumption was satisfied, a paired-samples t-test was chosen to analyze difference between pretest and post-test scores.

Quantitative Data Result based on Pretest and Post-Test

TABLE 4 | Paired Samples Statistics

		Mean	N	Std.	Std. Error
				Deviation	Mean
Pair 1	Pretest	75.75	8	1.282	.453
	Post-test	75.75	8	1.282	.453

Table 4 presents the descriptive statistics for the pretest and post-test debate simulation scores. Both tests produced an identical mean score of 75.75, indicating no noticeable improvement in participants' performance following the intervention. The standard deviation for both tests was 1.282, suggesting that the variability in scores remained consistent across the two stages. Similarly, the standard error of the mean was 0.453 in both tests, reflecting a comparable level of precision in estimating the sample mean.

In the pretest, students completed a debate without external assistance. In the post-test, ChatGPT was introduced as an intervention to support their argument development. The aim was to determine whether the integration of ChatGPT would enhance students' debate performance. However, the results showed no improvement, as reflected by the unchanged mean scores of 75.75 in both the pretest and post-test.

TABLE 5 | Paired Samples Correlations

		N	Correlation	Significance	
				One-Sided p	Two-Sided p
Pair 1	Pretest & Post-test	8	.826	.006	.011

Table 5 shows a fairly strong positive correlation ($r=0.826$) the pretest and post-test scores. This indicates that students who performed well in the pretest also tended to perform well in the post-test, while those who scored lower in the pretest showed similar outcomes in the post-test. In other words, students' debating skill remained relatively consistent across the two tests. The correlation was statistically significant ($p=0.006$ for the one-tailed test and $p = 0.011$ for the two-tailed test), suggesting that the relationship was unlikely to have occurred by chance.

However, despite this strong correlation, the mean score for both tests was identical (75.75). This indicates that the use of ChatGPT did not lead to any measurable improvement in

students' debating performance. In short, students' skills remained unchanged before and after the ChatGPT intervention.

TABLE 6 | Paired Samples Test

Pair		Paired Differences				t	df	Significance		
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference			One-Sided p	Two-Sided p	
					Lower					Upper
1	Pretest - Post-test	.000	.756	.267	-.632	.632	.000	7	.500	1.000

Table 6 shows the results of the paired samples t-test comparing students' scores in the pretest and post-test debate simulations. The t-value was 0.000, indicating no measurable difference between the two sets of scores. The two-tailed p-value was 1.000, which is far above the conventional threshold of 0.05 for statistical significance. This result confirms that the difference in scores between the pretest and post-test was not statistically significant.

TABLE 7 | Paired Samples Effect Sizes

Pair		Standardizer ^a	Point Estimate	95% Confidence Interval	
				Lower	Upper
1	Pretest - Cohen's d	.756	.000	-.693	.693
1	Post-test Hedges' correction	.851	.000	-.615	.615

a. The denominator used in estimating the effect sizes.

Cohen's *d* uses the sample standard deviation of the mean difference, while Hedges' correction applies the sample standard deviation with an added correction factor.

Table 7 presents the effect size statistics for the pretest and post-test scores. The Cohen's *d* value was 0.000, and the Hedges' correction was also 0.000. Both values fall within the 95% confidence interval, which ranged from -0.693 to 0.693 for Cohen's *d*, and -0.615 to 0.615 for Hedges' correction. Since these intervals include zero, the true effect could be negative, positive, or essentially nonexistent.

In other words, there was no measurable effect of the ChatGPT intervention on students' debate performance. Normally, Cohen's *d* values of 0.2, 0.5, and 0.8 are interpreted as small, medium, and large effects, respectively. In this study, the value of 0.000 shows the complete absence of an effect. The identical mean scores in the pretest and post-test (75.75) further confirm this result, showing that ChatGPT neither improved nor reduced performance.

At the same time, the relatively wide confidence intervals show the small sample size ($n = 8$), which makes the results less certain. This means that a larger study might reveal a small positive or negative effect. However, based on the current findings, it is reasonable to conclude that ChatGPT made no real difference to students' debating skills in this context.

The finding is important as the purpose of the study was to test whether ChatGPT could influence debate performance, and the results clearly showed that under the given conditions, it did not. The effect size outcome reflects how students interacted with ChatGPT: AI cannot automatically improve debating skills, especially when students rely on copying its output. Genuine improvement may require more training sessions, stronger guidance on how to craft prompts, and greater encouragement to critically engage with AI-generated responses.

Qualitative Data Results based on Semi-Structured Interviews

The semi-structured interviews focused on two main questions: (1) what kinds of information did the students seek when using ChatGPT? (2) How did they use and integrate the information provided by ChatGPT? The codes, categories, and themes that emerged from their responses are summarized in Table 8 below. These two questions were deliberately chosen to capture both the input (what information students looked for) and the processing (how they integrated and applied the information) aspects of AI-supported argumentation in debate. The first question highlights how students searched for information, while the second focuses the strategies they used to combine and make sense of it. This approach makes it possible to examine not only the material students obtained from ChatGPT but also how they absorbed and applied it in practice.

The thematic coding (see Table 8) revealed two overarching themes: matter in debate and integration of information. The first theme consisted of two categories: Debate strategy, which reflected students' responses about how they interpreted the motion and defined their team identity, and Argumentation Process, which included how students constructed arguments, added extensions, and supported them with relevant examples. The second theme, Integration of Information, demonstrated whether students engaged critically with ChatGPT's output (e.g., elaborating on their own ideas, consulting other sources, or constructing rebuttals) or relied on it more passively (e.g., copying arguments or depending too heavily on AI suggestions).

TABLE 8 | Codes, categories, and themes from students' responses

Code	Category	Theme
Interpretation of the motion	Debate Strategy	Matter in Debate
Identity of a team		
Argument	Argumentation Process	Integration of Information
Extension		
Example of argument	No reliance	Integration of Information
Self-elaboration		
Making self-rebuttals	AI reliance	Integration of Information
Using other credible sources		
ChatGPT suggestion	AI reliance	Integration of Information
Copying argument		

(1) What type of information did you look for when using ChatGPT?

PM: "I was looking for the interpretation of the motion. Even though it was easy to understand, I just wanted to make sure the meaning of "national exam". It turned out to be focusing on the score rather than the process." (motion interpretation)

DPM: "I was trying to find the possible arguments. I don't want to sound repetitive from my first speaker. Extension is quite hard to find. At least I get the main idea." (argument development)

LO: "I don't know what to ask at first, but I finally ask ChatGPT regarding a strategic identity for opening opposition." (motion interpretation)

DLO: "I am always not good at preventing opponent's argument. So, I really want to predict it and make preventive arguments, especially for my next opponent, Closing Government." (argument development)

MG: "As closing government, I want to simulate the opening debate. I asked ChatGPT the arguments both government and opposition. Then, I made alternative arguments to differentiate." (argument development)

GW: "The power of a whip relies on scenarios and examples in real life context. I am trying to predict the best and worst scenario of the motion, particularly in opposition team." (argument development)

MO: "I just asked ChatGPT to provide arguments for opposition. I think I need to anticipate it by asking the extension or the closing opposition arguments. So, I'm not trapped and repetitive." (argument development)

OW: "I am just helping my member to find arguments. I should not bring new arguments." (argument development)

(2) How did you use and integrate the information provided by ChatGPT?

PM: "The parameter given by ChatGPT makes me easier to contextualize the motion better. However, I still use my own description to elaborate the motion." (self-elaboration)

DPM: "I used the suggestions from ChatGPT on the main idea. I elaborate the points by myself." (self-elaboration)

LO: "I take the answer from ChatGPT. So, my team's identity was made by ChatGPT. It is hard for me because I don't have much time to understand it and immediately arrange my arguments." (ChatGPT elaboration)

DLO: "When I simulate the opponent's argument, especially Closing Government, I immediately make the rebuttals and modify it into my extension. I do not use ChatGPT to find my rebuttals, I like to search the information in credible sources. Yet, it helps me." (self-elaboration)

MG: "As prediction matters for Closing Government, I make the identity of Closing Government from the common answers of ChatGPT. It is faster as previously I need to memorize what arguments are possibly used. Now, I can elaborate my own argument faster. I don't use the extension from ChatGPT because it is still too common." (self-elaboration)

GW: "I do not take any suggestion from ChatGPT. I need to focus on my opponents' argument." (self-elaboration)

MO: "I almost copy all the arguments from ChatGPT. That's why I am a bit surprised when Opening Opposition has brought it. I find it hard to make a new identity for Closing Opposition." (ChatGPT elaboration)

OW: "I am taking some examples from ChatGPT for my case. I think it is still not enough. It is hard to see the case deeper only from the examples." (ChatGPT elaboration)

Theme 1: Matter in Debate

This theme shows how students used ChatGPT to engage directly with debate content. Two main categories emerged: *debate strategy* and *argumentation process*.

- Debate strategy included codes such as *motion interpretation* and *team identity*. Some students relied on ChatGPT to verify their understanding of the motion or to establish the strategic identity of their team. For example, the Prime Minister (PM) said, "I was looking for the interpretation of the motion. Even though it was easy to understand, I just wanted to make sure the meaning of 'national exam'." Similarly, the Leader of the Opposition (LO) reported using ChatGPT to shape the identity for the opposition bench.
- Argumentation process included *argument development*, *extensions*, and *examples*. Many students turned to ChatGPT for possible arguments or extensions to avoid repetition. For instance, the Deputy Prime Minister (DPM) admitted, "Extension is quite hard to find. At least I get the main idea." Meanwhile, the Government Whip (GW) focused on finding scenarios and real-life examples to strengthen their speech, while members of the opposition (MO, OW) also drew on ChatGPT for arguments or supporting examples.

Theme 2: Integration of Information

This theme highlights how students processed and applied the output provided by ChatGPT. Two distinct patterns emerged, *no reliance* and *AI reliance*.

- *No reliance* referred to students who critically engaged with ChatGPT's suggestions, often modifying them or cross-checking with other sources. For example, the Deputy Leader of the Opposition (DLO) shared, "I immediately make the rebuttals and modify it into my extension. I do not use ChatGPT to find my rebuttals, I like to search the information in credible sources." Similarly, the Closing Government Member (MG) explained that while ChatGPT saved time in predicting common arguments, they still developed their own unique points.
- *AI reliance* described students who depended heavily on ChatGPT without much critical engagement. For example, the LO admitted that their team identity was taken directly from ChatGPT, saying, "It is hard for me because I don't have much time to understand it." Likewise, the MO confessed, "I almost copy all the arguments from ChatGPT. That's why I am a bit surprised when Opening Opposition has brought it."

The interview findings show that students used ChatGPT primarily to understand concepts and build arguments. While all participants actively engaged with the tool during the post-test debate simulations, their approaches varied. Some treated ChatGPT's suggestions as starting points, strengthening their arguments by adding personal insights and cross-checking with reliable sources. Others, however, relied almost entirely on ChatGPT's direct responses, seeking ready-made arguments to support their positions.

A clear performance gap emerged between these two groups. Students who relied too heavily on ChatGPT showed little to no improvement in their post-test scores, and in some cases, their scores declined compared to the pretest. This likely occurred because they tended to copy ChatGPT's responses without fully developing their own arguments. By contrast, students who used ChatGPT more critically and strategically demonstrated stable or slightly improved performance, reflecting deeper engagement with the material.

These findings align with the results of the paired sample t-test, which indicated minimal improvement in overall post-test performance. This suggests that while ChatGPT provided accessible information, many students lacked effective strategy to critically evaluate and integrate AI-generated content, leading to over-reliance and limited analytical engagement.

This study investigates the impact of AI-generated responses, specifically from ChatGPT, on students' debate performance. It examines how students used ChatGPT to build cases and gathering information to support their arguments. The findings indicate that ChatGPT provided clear support in retrieving information, which helped students organize their debate points more effectively. This aligns with [Ba et al. \(2025\)](#), who reported that ChatGPT-assisted feedback in online learning discussions improved students' ability to develop and structure their reasoning. Similarly, students in this study used ChatGPT to clarify motions, shape arguments, and identify relevant examples.

Despite these advantages, the study found that ChatGPT did not significantly improve students' debate performance. Although a moderate correlation appeared between pretest and post-test scores, the paired samples t-test showed no

statistically meaningful difference ($p = 1.000$). The effect size (Cohen's $d = 0.000$) further confirmed that the intervention had no measurable impact. In other words, the small changes observed in students' score were most likely due to chance rather than the direct influence of ChatGPT. This outcome contrasts with earlier studies that reported positive effects of ChatGPT and similar AI tools on students' speaking and debating skills ([Alanzi et al., 2023](#); [Almazayad et al., 2023](#); [Karakose & Tülübaş, 2023](#)). However, it aligns with research by [Han et al. \(2022\)](#), which also found no significant improvement in reasoning skills from AI support. These differences suggest that the effectiveness of ChatGPT depends on the learning context, the duration of the intervention, and the extent of guidance students receive in using AI for academic tasks.

Although ChatGPT did not lead to significant gains in debate scores, it appeared to support growth in critical thinking. Interview data revealed that students often treated ChatGPT's responses as prompts to develop their own arguments, showing evidence of analysis and reflection. This echoes findings from other studies ([Minh, 2024](#); [Ruff et al., 2024](#); [Xiao & Zhi, 2023](#)), which showed that students enhanced their critical and analytical thinking by recognizing the limitations of AI-generated content and expanding upon it independently.

A key reason for the lack of score improvement was students' over-reliance on ChatGPT. Interview data indicated that some students copied answers directly from ChatGPT or failed to critically assess the information, resulting in weaker arguments and, in some cases, lower post-test scores. This finding echoes previous research ([Guleria et al., 2023](#); [Petrić, 2024](#); [Shehri et al., 2023](#)), which shows that excessive dependence on AI can hinder critical thinking by bypassing deeper research and synthesis.

In addition to over-reliance, students faced challenges in using ChatGPT effectively. Many felt confused when they received repetitive answers, often similar to those given to other groups, and struggled to create prompts that could generate diverse and useful responses. [Nugroho \(2023\)](#) also highlighted that students risk misunderstanding or being misled by AI when they fail to critically evaluate its outputs. Supporting this, several studies ([Liu, 2023](#); [Liu et al., 2023](#); [Miyazaki, 2024](#); [Rashid et al., 2024](#); [Seniwati et al., 2024](#)) have found that ChatGPT responses, while often appearing credible, can be inaccurate or misleading. The findings underscore the need for clearer strategies and stronger guidance to ensure students rely on trustworthy information when using ChatGPT or similar AI tools.

Overall, these findings highlight the important role of educators in guiding students' engagement with AI-generated content. Teachers need to help students avoid pitfalls such as over-reliance, confusion, and misinformation while fostering critical engagement with AI outputs. [Đoàn \(2023\)](#) emphasized that educators can reduce students' cognitive load, enabling more effective learning. Therefore, educators should not only introduce students to ChatGPT but also train them to critically analyze and integrate AI-generated information. Additionally, structured curricula that thoughtfully incorporate AI tools are urgently needed to ensure their use enhances, rather than detracts from, learning

outcomes.

It is also important to recognize the limitations of this study. The sample size was very small ($n = 8$), which limits how much the findings can be generalized. The design was a one-group pre-experimental study without a control group, limiting the ability to isolate the effects of ChatGPT from other factors. The treatment period was short (three meetings), which may not have been sufficient for students to develop stronger debate strategies or learn to use ChatGPT effectively. Limitations also arose from ChatGPT itself: its responses were sometimes generic or repetitive, and students with weaker prompting skills struggled to refine them. Finally, individual variation likely played a role. Some students, especially those who were more confident, tended to use ChatGPT critically, while others relied on it passively.

CONCLUSION

This study examines the impact of AI-generated responses, specifically from ChatGPT, on the debate performance of university students. The findings indicate that using ChatGPT use did not lead to any significant improvement in students' debating skills. Despite a strong correlation between pretest and post-test scores, the identical mean score of 75.75 for both assessment suggests that the intervention had no measurable effect. This consistency implies that students' abilities remained stable throughout the study, with ChatGPT failing to produce any substantial change. The t-test results ($p = 1.000$) further support this conclusion, indicating that any observed differences in performance were likely due to chance rather than the influence of ChatGPT. In relation to the research question, the study concludes that ChatGPT, when used as a feedback tool in a short intervention does not significantly enhance students' debate performance. Future research could explore alternative strategies or investigate other aspects of debate simulation to better support student learning.

The lack of performance is largely attributed to students limited critical engagement with AI-generated responses. Many struggled to evaluate the information critically, often copying or relying entirely on ChatGPT's content without deeper analysis. This underscores the vital role of educators in supporting students' effective use of AI tools. A key implication is that teachers should provide explicit instruction on how to engage critically with ChatGPT outputs, for example, by training students in effective prompting, encouraging them to cross-check AI responses with credible sources, and guiding them to elaborate on ideas independently. Such strategies could help prevent over-reliance and support deeper learning.

As AI use in education continues to expand, further research on how to best leverage AI-generated responses is crucial. This study has several limitations that must be considered, such as the small sample size ($n = 8$), the short duration of the intervention (three sessions), the absence of a control group, and the context-specific nature of debating, which may not generalize to other disciplines. These limitations suggest that the findings should be interpreted with caution. Nevertheless, they highlight opportunities for

future research to test AI-assisted debate practice with larger and more diverse samples, over longer time frames, and with comparisons to traditional methods.

In conclusion, while this study found no measurable effect of ChatGPT on debate performance, it provides valuable insights into how students interact with AI tools, sometimes engaging with them critically, but at other times relying on them passively. This suggests that more structured guidance is needed to help students maximize the educational benefits of AI. Future studies should not only explore the long-term impacts of AI-based feedback but also examine how sustained training with AI can foster critical thinking and assess its effectiveness across different disciplines and educational contexts.

ACKNOWLEDGEMENTS

The researchers sincerely thank the volunteers who participated in this study. They are also grateful for the valuable and constructive feedback provided by the journal's reviewers, which greatly improved the quality of this paper.

REFERENCES

- Abbas, N., Imran, A., Manzoor, R., Hussain, T., & Hussain, M. (2023). Role of artificial intelligence tools in enhancing students' educational performance at higher levels. *Journal of Artificial Intelligence Machine Learning and Neural Network*, (35), 36-49. <https://doi.org/10.55529/jaimlnn.35.36.49>
- Aisyah, Yulianti, P.D., Yandhini, S., Sari, A.D.P., Herawani, I., & Oktarini, Ira. (2024). The Influence of AI on Students' Mind Patterns. *BICC Proceedings*, 2, 183-186. <https://doi.org/10.30983/bicc.v1i1.125>
- Akopiants, N. (2023). Using ChatGPT in The Process of Learning English: Advantages and Opportunities. *Bulletin of the National Technical University "KhPI". Series: Actual Problems of Ukrainian Society Development*, 1, 69–72. <https://doi.org/10.20998/2227-6890.2023.1.13>
- Alanzi, T., Alzahrani, W., Albalawi, N., Allahyani, T., Alghamdi, A., Al-Zahrani, H., & Farhah, N. (2023). Public awareness of obesity as a risk factor for cancer in central Saudi Arabia: Feasibility of ChatGPT as an educational intervention. *Cureus*. <https://doi.org/10.7759/cureus.50781>
- Alasmari, A. & Ahmed, S. (2012). Using debate in EFL classes. *English Language Teaching*, 6(1). <https://doi.org/10.5539/elt.v6n1p147>
- Ali, M., Faridi, A., & Sakhiyya, Z. (2020). Performance of Critical Thinking and Existence of Logical Fallacies in Indonesian Varsity English Debate 2020 in Jakarta. *International Journal of Research and Studies Publishing*, 11(1), 655-659. <https://doi.org/10.29322/ijrsrp.11.01.2021.p10980>
- Almazyad, M., Aljofan, F., Abouammoh, N. A., Muaygil, R., Malki, K. H., Aljamaan, F., Alhaboob, A., Alhasan, K., Alabdulkareem, A., Alhussien, N., Aloriny, I.,

- Alshehri, A., Alzahrani, S., Alyazidi, A., Alkhateeb, A., Alfaraaj, D., Abokhodair, N., Almuzaini, Y., & Temsah, M.-H. (2023). Enhancing expert panel discussions in pediatric palliative care: Innovative scenario development and summarization with ChatGPT-4. *Cureus*, 15(4), e38249. <https://doi.org/10.7759/cureus.38249>
- Astuti, A., Suryani, F., & Kurniati, D. (2022). The EFL student teachers' debate practice: problems and solutions. *Prominent*, 5(1), 72-78. <https://doi.org/10.24176/pro.v5i1.7028>
- Ba, S., Zhan, Y., Huang, L., & Lu, G. (2025). Investigating the impact of ChatGPT-assisted feedback on the dynamics and outcomes of online inquiry-based discussion. *British Journal of Educational Technology*. *Advance online publication*. <https://doi.org/10.1111/bjet.13605>
- Ban, B., Pang, S., & Em, S. (2023). Debate: one of the key factors to improving students' English language speaking skills. *Journal of General Education and Humanities*, 2(2), 107-120. <https://doi.org/10.58421/gehu.v2i2.69>
- Barabad, M.G., & Anwar, B.M. (2024). Exploring ChatGPT's Role in English Learning for EFL Students: Insights and Experiences. *International Journal of Innovative Science and Research Technology (IJISRT)*, 755-766. <https://doi.org/10.38124/ijisrt/ijisrt24sep347>
- Baziganya, D. (2024). Investigating teachers and students' views on debate activities as a tool to improve English speaking skills: a case of Rulindo district, rwanda. *African Journal of Empirical Research*, 5(2), 869-880. <https://doi.org/10.51867/ajernet.5.2.77>
- Chen, X., Wang, L., Zhai, X., & Li, Y. (2022). Exploring the effects of argument map-supported online group debate activities on college students' critical thinking. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.856462>
- Chew, Q., Seet, X., & Sim, K. (2021). Use of debate as a pedagogical tool in psychiatry residency teaching: a cross-sectional study. *Advances in Medical Education and Practice*, 12, 871-877. <https://doi.org/10.2147/amep.s320614>
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: qualitative, quantitative, and mixed methods approaches (Fifth edition)*. Los Angeles: SAGE.
- Daniswara, L. D., & Cahyono, B.Y. (2023). Investigation of university students' critical thinking in debate: Justification for the "AREL" argumentation process. *JEES (Journal of English Educators Society)*, 8(2). <https://doi.org/10.21070/jees.v8i2.1766>
- Didonna, N., Shetty, P. N., Khan, K., & Damitz, L. (2024). Unveiling the Potential of AI in Plastic Surgery Education: A Comparative Study of Leading AI Platforms' Performance on In-training Examinations. Plastic and Reconstructive Surgery. *Global Open*, 12(6). <https://doi.org/10.1097/gox.0000000000005929>
- Đoàn, T.N. (2023). Motivation from Technology Innovation in Teaching Foreign Languages: A Case Study in Vietnam. *Vietnam Academy of Social Sciences's VSSR*, 214(2), 92-104. [https://doi.org/10.56794/VSSR.2\(214\).92-104](https://doi.org/10.56794/VSSR.2(214).92-104)
- Esmail, A., Dzulkipli, D., Maakip, I., Mantaluk, O., & Marshall, S. (2023). Understanding student perception regarding the use of ChatGPT in their argumentative writing: a qualitative inquiry. *Jurnal Komunikasi Malaysian Journal of Communication*, 39(4), 150-165. <https://doi.org/10.17576/jkmjc-2023-3904-08>
- Guleria, A., Krishan, K., Sharma, V., & Kanchan, T. (2023). Chatgpt: ethical concerns and challenges in academics and research. *The Journal of Infection in Developing Countries*, 17(09), 1292-1299. <https://doi.org/10.3855/jidc.18738>
- Gupta, D., Sreelatha, D., Latha, A., Raj, D., & Singh, D. (2024). Navigating the future of education: the impact of artificial intelligence on teacher-student dynamics. *EATP*, 6006-6013. <https://doi.org/10.53555/kuvey.v30i4.2332>
- Halaweh, M. (2023). ChatGPT in education: Strategies for responsible implementation. *Contemporary Educational Technology*, 15(2), ep421. <https://doi.org/10.30935/cedtech/13036>
- Han, J., Park, J., & Lee, H. (2022). Analysis of the effect of an artificial intelligence chatbot educational program on non-face-to-face classes: a quasi-experimental study. *BMC Medical Education*, 22(1). <https://doi.org/10.1186/s12909-022-03898-3>
- Hernawan, A. & Yusuf, F. (2021). English competitive debating challenges. Atlantis Press, *Proceedings of the Thirteenth Conference on Applied Linguistics (CONAPLIN 2020)*. <https://doi.org/10.2991/assehr.k.210427.007>
- Idroes, G. M., Novandy, T. R., Maulana, A., Irvanizam, I., Jalil, Z., Lensoni, L., Lala, A., Abas, A. H., Tallei, T. E., & Idroes, R. (2023). Student perspectives on the role of artificial intelligence in education: A survey-based analysis. *Journal of Educational Management and Learning*, 1(1), 8-15. <https://doi.org/10.60084/jeml.v1i1.58>
- Ifraheem, S., Rasheed, M., & Siddiqui, A. (2024). Transforming education through artificial intelligence: personalization, engagement and predictive analytics. *J. Asian Dev. Studies*, 13(2), 250-266. <https://doi.org/10.62345/jads.2024.13.2.22>
- Ivanova, M. & Petrova, T. (2023). Towards independent students' activities, online environment and learning performance: an investigation through synthetic data and artificial neural networks. *MDPI*, 10(2), 37. <https://doi.org/10.3390/informatics10020037>
- Jiang, F. & Hyland, K. (2024). Does ChatGPT Argue Like Students? Bundles in Argumentative Essays, *Applied Linguistics*, amae052. <https://doi.org/10.1093/applin/amae052>
- Karakose, T. & Tülübaş, T. (2023). How can ChatGPT facilitate teaching and learning: implications for contemporary education. *Educational Process International Journal*, 12(4). <https://doi.org/10.22521/edupij.2023.124.1>

- Ko, T. & Mezuk, B. (2021). Debate participation and academic achievement among high school students in the Houston independent school district: 2012 - 2015. *Educational Research and Reviews*, 16(6), 219-225. <https://doi.org/10.5897/err2021.4137>
- Lestari, Y. (2018). Students' perceptions toward speaking achievement and critical thinking on the use of British parliamentary debating system. *Els Journal on Interdisciplinary Studies in Humanities*, 1(4), 441-447. <https://doi.org/10.34050/els-jish.v1i4.5187>
- Li, M., & Wilson, J. (2025). AI-integrated scaffolding to enhance agency and creativity in K-12 English Language Learners: A systematic review. *Information*, 16(7), Article 519. <https://doi.org/10.3390/info16070519>
- Liu, C.-L., Ho, C.-T., & Wu, T.-C. (2024). Custom GPTs Enhancing Performance and Evidence Compared with GPT-3.5, GPT-4, and GPT-4o? A Study on the Emergency Medicine Specialist Examination. *Healthcare (Basel, Switzerland)*, 12(17). <https://doi.org/10.3390/healthcare12171726>
- Liu, H., Azam, M., Naeem, S., & Faiola, A. (2023). An overview of the capabilities of ChatGPT for medical writing and its implications for academic integrity. *Health Information & Libraries Journal*, 40(4), 440-446. <https://doi.org/10.1111/hir.12509>
- Liu, Z. (2023). Research on the impact of ChatGPT on the accounting industry and countermeasures. *Frontiers in Business, Economics and Management*, 11(2), 49-53. <https://doi.org/10.54097/fbem.v11i2.12162>
- Minh, A. (2024). Leveraging ChatGPT for enhancing English writing skills and critical thinking in university freshmen. *Journal of Knowledge Learning and Science Technology ISSN 2959-6386 (Online)*, 3(2), 51-62. <https://doi.org/10.60087/jklst.vol3.n2.p62>
- Miyazaki, Y. (2024). Performance and errors of chatgpt-4o on the Japanese medical licensing examination: solving all questions including images with over 90% accuracy. *JMIR Medical Education*. <https://doi.org/10.2196/preprints.63129>
- Munir, H., Vogel, B., & Jacobsson, A. (2022). Artificial intelligence and machine learning approaches in digital education: a systematic revision. *MDPI*, 13(4), 203. <https://doi.org/10.3390/info13040203>
- Nguyen, H., Ngoc, H., & Dan, T. (2024). EFL students' perceptions and practices of using ChatGPT for developing English argumentative essay writing skills. *European Journal of Alternative Education Studies*, 9(1). <https://doi.org/10.46827/ejae.v9i1.5341>
- Nugroho, A., Syamsi, K., & Putro, N. H. P. S. (2023). The Potentials of ChatGPT for Language Learning: Unpacking its Benefits and Limitations. *Register Journal*, 16(2), 224-247. <https://doi.org/10.18326/register.v16i2.224-247>
- Nurakhir, A., Palupi, F., Langeveld, C., & Nurmalia, D. (2020). Students' views of classroom debates as a strategy to enhance critical thinking and oral communication skills. *Nurse Media Journal of Nursing*, 10(2), 130-145. <https://doi.org/10.14710/nmjn.v10i2.29864>
- Onesi-Ozigagun, O. (2024). Revolutionizing education through ai: a comprehensive review of enhancing learning experiences. *International Journal of Applied Research in Social Sciences*, 6(4), 589-607. <https://doi.org/10.51594/ijarss.v6i4.1011>
- Oon, M., Syn, N., Tan, N., Tan, K., & Ng, S. (2023). Bridging bytes and biopsies: a comparative analysis of ChatGPT and histopathologists in pathology diagnosis and collaborative potential. *Histopathology*, 84(4), 601-613. <https://doi.org/10.1111/his.15100>
- Pahmi, P. & Niah, S. (2021). Students' voices towards online debate through WhatsApp as alternative media to enhance critical thinking skills during covid-19 pandemic. *International Journal of Interactive Mobile Technologies (IJIM)*, 15(05), 205. <https://doi.org/10.3991/ijim.v15i05.18759>
- Pang, S., Nol, E., & Heng, K. (2024). ChatGPT-4o for English language teaching and learning: Features, applications, and future prospects. *Cambodian Education Forum*, (4)1, 35-56. <http://dx.doi.org/10.2139/ssrn.4837988>
- Petrič, G. (2024). Everyone talks everything with ChatGPT: Students' uses of ChatGPT and their impact on learning performance. *International Journal of Technology and Human Interaction*, 20(1), 1-22. <https://doi.org/10.4018/ijthi.349225>
- Prompiengchai, P., Narreddy, P., & Joordens, S. (2025). A Practical Guide for Supporting Formative Assessment and Feedback Using Generative AI. <https://doi.org/10.48550/arXiv.2505.23405>
- Rahmi, A. (2021). High school debate: elevating students critical thinking into hoax prevention. Atlantis Press, *Proceedings of the 2nd Southeast Asian Academic Forum on Sustainable Development (SEA-AFSID 2018)*. <https://doi.org/10.2991/aebrm.k.210305.069>
- Rashid, A., Skelly, R., Valdes, C., Patel, P., Solberg, L., Giordano, C., Modave, F. (2024). Evaluating chatgpt's moral competence in health care-related ethical problems. *Jamia Open*, 7(3). <https://doi.org/10.1093/jamiaopen/ooae065>
- Ruff, E., Engen, M., Franz, J., Mauser, J., West, J., & Zemke, J. (2024). Chatgpt writing assistance and evaluation assignments across the chemistry curriculum. *Journal of Chemical Education*, 101(6), 2483-2492. <https://doi.org/10.1021/acs.jchemed.4c00248>
- Ruff, E. F., Franz, J. L., & West, J. K. (2024). Using chatgpt for method development and green chemistry education in upper-level laboratory courses. *Journal of Chemical Education*, 101(8), 3224-3232. <https://doi.org/10.1021/acs.jchemed.4c00193>
- Sanonguthai, S. (2011). Teaching IELTS writing module through English debate: a case study in Thailand. *Language Testing in Asia*, 1(4). <https://doi.org/10.1186/2229-0443-1-4-39>
- Sawyer, D. (2015). Step your game up. *Humanity & Society*, 39(2), 224-235. <https://doi.org/10.1177/0160597615574552>
- Sen, T. K., Naven, G., Gerstner, L., Bagley, D., Baten, R. A., Rahman, W., Hasan, M. K., Haut, K. G., Mamun, A., Samrose, S., Solbu, A., Barnes, R. E., Frank, M. G., &

- Hoque, E. (2021). DBATES: Database of Audio features, Text, and visual Expressions in competitive debate Speeches. *IEEE Transactions on Affective Computing*, 14(2), 1028–1043. <https://doi.org/10.1109/TAFFC.2021.3103442>
- Seniwati, S., Lumumba, P., & Culla, A. S. (2024). The effect of GPT chat on Indonesian and Italian national security. *International Journal of Social Science Humanity & Management Research*, 3(01). <https://doi.org/10.58806/ijsshr.2024.v3i1n08>
- Shehri, F. A., Maham, R., Malik, A., & Saif, O. B. (2023). Effects of ChatGPT on students' academic performance: Mediating role of prompt engineering. *The Asian Bulletin of Big Data Management*, 3(2), 137-147. <https://doi.org/10.62019/abbdm.v3i2.58>
- Spaska, A., Savishchenko, V., Komar, O., Hritchenko, T., & Maidanyk, O. (2021). Enhancing analytical thinking in tertiary students using debates. *European Journal of Educational Research*, 10(2), 879-889. <https://doi.org/10.12973/eu-jer.10.2.879>
- Sugianti & Rosidah, I. (2024). The use of artificial intelligence (ai) in learning results for scientific Indonesian language courses at PGRI Wiranegara university. *IJARSS*, 2(1), 1-10. <https://doi.org/10.59890/ijarss.v2i1.1174>
- Tjahyaningtjas, H. (2023). Machine learning on academic education: bibliometric studies. *E3s Web of Conferences*, 450, 02010. <https://doi.org/10.1051/e3sconf/202345002010>
- Trisnawati, W. (2023). The impact of artificial intelligent in education toward 21st century skills: a literature review. *PIJED*, 2(2), 501-513. <https://doi.org/10.59175/pijed.v2i2.152>
- Wang, B., Sun, H., & Yue, X. (2023). Can ChatGPT Defend its Belief in Truth? Evaluating LLM Reasoning via Debate. *Cornell university*. <https://doi.org/10.48550/arxiv.2305.13160>
- Wang, Y. (2024). Reviewing the usage of ChatGPT on 12 students' English academic writing learning. *Journal of Education Humanities and Social Sciences*, 30, 173-178. <https://doi.org/10.54097/dvjki706>
- Wariyati, W. (2023). Overcoming speaking barrier. *Journey (Journal of English Language and Pedagogy)*, 6(3), 713-720. <https://doi.org/10.33503/journey.v6i3.3775>
- Xiao, Y. & Zhi, Y. (2023). An exploratory study of EFL learners' use of ChatGPT for language learning tasks: experience and perceptions. *Languages*, 8(3), 212. <https://doi.org/10.3390/languages8030212>
- Xu, J. (2023). Students' cognition of artificial intelligence in education. *Applied and Computational Engineering*, 22(1), 200-209. <https://doi.org/10.54254/2755-2721/22/20231218>
- Yang, C., Chiang, F., Cheng, Q., & Ji, J. (2021). Machine learning-based student modeling methodology for intelligent tutoring systems. *Journal of Educational Computing Research*, 59(6), 1015-1035. <https://doi.org/10.1177/0735633120986256>

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright ©2025 Landry Dwiyoga Daniswara. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Implementing and evaluating OpenAI whisper for accurate speaking assessment and skill development in Indonesian EFL classroom

Salsabila Latifa*, Nirwanto Maruf

Universitas Muhammadiyah Gresik, Indonesia

This study investigates the implementation and effectiveness of an OpenAI Whisper-based automatic speech recognition (ASR) system for evaluating and improving the speaking skills of Indonesian EFL students. Employing a mixed method, one group pretest–posttest design, the research involved 40 undergraduate participants. Quantitative data were collected through standardized speaking tests rated by both the Whisper system and expert human assessors, focusing on fluency, pronunciation, and coherence. Qualitative insights were obtained from classroom observations and in depth interviews with students and lecturers, exploring user experiences and contextual factors affecting system performance. The results demonstrate that the Whisper based assessment system achieved high inter-rater reliability with human experts (Cohen’s Kappa = 0.81; ICC = 0.87) and led to significant improvements in learners’ speaking skills across all assessed dimensions. Implementation of the Whisper based intervention produced statistically significant pre–post gains (all $p < .001$) with large effect sizes: overall performance ($d = 1.02$), fluency ($d = 0.97$), pronunciation ($d = 1.11$), and coherence ($d = 1.00$). The system’s immediate, actionable feedback fostered greater learner engagement and autonomy, with pronunciation showing the largest gains. However, technological infrastructure, digital literacy, and classroom conditions influenced the intervention’s effectiveness and reliability. These findings highlight the importance for robust infrastructure, teacher training, and equitable access to technology. The study validates a multidimensional, context adaptive framework for AI based speaking evaluation and offers practical guidelines for integrating ASR into EFL curricula, urging educators and policymakers to prioritize funding for infrastructure, teacher professional development, and digital literacy programs.

Keywords: OpenAI Whisper, Automatic Speech Recognition (ASR), Speaking Skills, AI-Driven Assessment, Digital Literacy

OPEN ACCESS
ISSN 2503 3492 (online)

*Correspondence:
Salsabila Latifa

salsabilatfa1210@gmail.com

Received: 1st July 2025

Accepted: 20th September 2025

Published: 26th October 2025

Citation:

Latifa, S., Maruf, N. (2025). Implementing and evaluating OpenAI whisper for accurate speaking assessment and skill development in Indonesian EFL classroom. *JEES (Journal of English Educators Society)*, 10(2). <https://doi.org/10.21070/jees.v10i2.1952>

INTRODUCTION

Despite decades of innovation in English as a Foreign Language (EFL) education, only 30% of Indonesian university students achieve the minimum proficiency in spoken English required for academic and professional contexts (Fajrina et al., 2021; Maruf et al., 2020). This persistent gap is not merely a matter of language exposure or curriculum design, it is deeply tied to how speaking skills are assessed in classroom settings (Coleman et al., 2024; Irham et al., 2022; Munandar & Shaumiwyaty, 2023; Prasandha & Aniq, 2023). Traditional assessment methods, which rely heavily on human raters, are often criticized for their subjectivity and inconsistency,

with studies indicating that up to 35% of scoring variance can be attributed to rater bias and fatigue ([Henze et al., 2024](#); [Isaacs & Thomson, 2013](#)). Consequently, learners often receive feedback that is not only inconsistent but also insufficiently actionable, limiting both their motivation and their progress ([Alfredo et al., 2024](#); [Geva, 2017](#); [Kahng, 2023](#); [Winke & Gass, 2013](#)).

Recent advances in educational technology have prompted growing interest in leveraging artificial intelligence (AI) to address these limitations. Automatic Speech Recognition (ASR) systems, such as OpenAI Whisper, have demonstrated significant potential in transcribing and evaluating non-native English speech with high accuracy, reportedly achieving transcription rates above 90% for diverse accents ([Alharbi et al., 2021](#); [Bhardwaj et al., 2022](#); [Dhouib et al., 2022](#); [Fendji et al., 2022](#)). However, the practical integration of these technologies into EFL classrooms remains limited, particularly in contexts with unique linguistic features like Indonesia ([Alharbi et al., 2021](#); [Feng et al., 2024](#); [Santhanavijayan et al., 2021](#); [Wu et al., 2023](#)).

The core challenge, therefore, is the lack of an objective, scalable, and context-sensitive system for assessing EFL speaking skills, particularly in countries like Indonesia, where linguistic diversity and large class sizes complicate reliable evaluation. While AI and ASR technologies have shown considerable promise, most existing systems focus narrowly on transcription or isolated dimension such as pronunciation, without addressing fluency and coherence as integrated components ([Cengiz, 2023](#); [Malik et al., 2021](#); [Yuan & Liu, 2020](#); [Jiang et al., 2021](#)). Moreover, the applicability of these technologies in real classroom settings, especially with non-native accents and local linguistic features, remains underexplored ([McGuire, 2025](#); [Ding et al., 2022](#); [Tejedor-Garcia et al., 2020](#)).

Several recent studies have explored the integration of ASR technologies in EFL speaking assessment, yet their approaches and findings reveal important gaps that the present research seeks to address. For example, [Bashori et al. \(2024\)](#) investigated two ASR-based language learning systems, ILI and NovoLearning, among Indonesian EFL learners. Their study found that both systems significantly improved students' English pronunciation at the word and sentence levels, as measured by phonetic edit distance, degree of accent, and comprehensibility. Notably, the NOVO system, which provides detailed phonetic feedback, led to greater improvements than ILI, which offers only global corrective feedback. However, while these ASR tools proved effective for pronunciation, their evaluation did not fully encompass other critical aspects of speaking proficiency such as fluency and coherence, thus providing only a partial picture of students' overall speaking abilities.

Similarly, [McGuire \(2025\)](#) research demonstrates the feasibility and reliability of fully automated speaking tests using Whisper ASR combined with elicited imitation (EI) and Word Error Rate (WER) scoring, showing near-perfect alignment between automated scoring and human raters. His study highlights the scalability, accessibility, and cost-effectiveness of such systems for large-scale language proficiency assessment, emphasizing their potential for

frequent, real-time evaluation and the development of adaptive, curriculum-specific tests. However, McGuire's work focuses primarily on sentence-repetition tasks under controlled conditions, with less attention to spontaneous speech or holistic, multidimensional assessment of speaking skills beyond pronunciation accuracy and transcription reliability, leaving these areas underexplored.

In contrast, the current study not only implements OpenAI Whisper to evaluate fluency, pronunciation, and coherence simultaneously but also rigorously compares its performance with expert human raters. Additionally, it investigates contextual factors influencing system accuracy in Indonesian EFL classrooms. This multidimensional and context-sensitive approach addresses critical gaps in prior research, offering a more comprehensive understanding of both the potential and limitations of AI-driven assessment systems for speaking skills in diverse educational settings.

The current study aims to implement and rigorously evaluate the effectiveness of an OpenAI Whisper-based assessment system in improving both the accuracy of evaluation and the speaking skills of EFL students. Specifically, its objectives are: (1) to compare the accuracy of the Whisper-based system with expert human raters across fluency, pronunciation, and coherence; (2) to evaluate the system's impact on student speaking performance; and (3) to identify contextual factors that influence the system's effectiveness in Indonesian EFL classrooms. By addressing these objectives, this research contributes to both theory and practice. Theoretically, it extends the literature on AI-driven language assessment by validating a multidimensional, context-adaptive framework for speaking evaluation. Practically, it provides empirical evidence and implementation guidelines for integrating advanced ASR technology into EFL curricula, paving the way for more equitable, consistent, and actionable assessment practices in diverse educational contexts. Ultimately, the findings are expected to inform policymakers, educators, and technologists in developing scalable solutions tailored to the unique needs of Indonesian learners and comparable EFL settings.

Based on the study's objectives, the following research questions guide this investigation: (1) How does the accuracy of the OpenAI Whisper-based assessment system compare with expert human raters in evaluating fluency, pronunciation, and coherence of EFL students' speaking performances? (2) To what extent does the implementation of the Whisper-based system enhance the speaking skills of Indonesian EFL learners? (3) What contextual factors within Indonesian EFL classrooms influence the effectiveness and reliability of the OpenAI Whisper-based assessment system? Addressing these questions provides comprehensive insights into both the technical validity of AI-driven assessment and its practical applicability in real-world educational settings, thereby contributing valuable knowledge to the fields of language assessment and educational technology.

METHODS

Research Design

This study employed a pre-experimental, one-group pretest-

posttest design with a mixed-methods approach. The quantitative component measured changes in students' speaking skills before and after the intervention using standardized speaking tests, while the qualitative component explored participants' experiences and contextual factors through observations and interviews. This design was chosen to provide both objective measurement of learning outcomes and in-depth insights into implementation challenges, thereby enhancing the reliability and comprehensiveness of the findings.

Participants and Setting

The research involved 40 undergraduate students enrolled in the English Education Department at Universitas Muhammadiyah Gresik, Indonesia. Participants were selected according to the following inclusion criteria: (1) active enrollment in semesters 4-6, (2) a minimum intermediate English proficiency (TOEFL PBT ≥ 500), (3) willingness to participate throughout the study, and (4) no concurrent enrollment in external speaking courses. The study was conducted in the university's language laboratory, which was equipped with computers and audio devices compatible with the OpenAI Whisper system.

The study adhered to ethical standards for studies involving human participants and complied with the principles of the Declaration of Helsinki. The research protocol and consent procedures were reviewed and approved by the Research Ethics Committee of Universitas Muhammadiyah Gresik. All 40 participants provided written informed consent. Participation was voluntary, and students were informed of their right to withdraw at any time without penalty. Audio recordings, transcripts, and assessment data were anonymized and securely stored on password protected institutional servers accessible only to the research team. Identifiable consent forms were stored separately in accordance with university policy.

TABLE 1 | Demographic and Proficiency Profile of Participants

Characteristic	Statistic / Category	
Sample size	—	40
Age	Mean ± SD	20.8 ± 1.1 y
	Range	19–23 years old
Gender	Male	12
	Female	28
Academic semester	Semester 4	14 (35%)
	Semester 5	16 (40%)
	Semester 6	10 (25%)
TOEFL PBT score	Range	500-560
	500–519	18 (45%)
Proficiency level (TOEFL PBT)	(Intermediate)	
	520–539 (Upper-intermediate)	15 (37.5%)
	≥ 540 (Advanced)	7 (17.5%)

Research Procedures

The research was conducted in several stages, as outlined in [Table 2](#):

TABLE 2 | Research procedures

Stage	Main Procedure	Replication Note
Preliminary	Run Whisper on 10 sample recordings; check ASR accuracy	Whisper version, audio specifications, and acceptability threshold recorded
Recruitment	Enroll 40 Indonesian EFL students (TOEFL PBT ≥ 500) with signed consent	Sampling frame and consent form archived
Needs analysis	Five lecturer interviews and student questionnaire	Interview guide and questionnaire items documented
Preparation	Train lecturers; finalize fluency-pronunciation-coherence rubric	Rubric and training slides provided
Intervention	16 speaking sessions over 3 months; pre/post-tests scored by Whisper and two expert raters	Session plan, test prompts, and rating sheets available
Post-intervention	Post-test, classroom observations, 5 student and 2 lecturer interviews	Same rubric and observation checklist used
Analysis	Paired t-test, Cohen's κ, ICC (quantitative); thematic analysis (qualitative)	

Data Collection

Pilot Testing and Instrument Validation

A pilot test was conducted prior to the main intervention using 10 randomly selected student speaking samples. Its primary aim was to evaluate the technical accuracy and operational feasibility of the OpenAI Whisper system in transcribing and scoring non-native English speech within the Indonesian EFL context. During the pilot, both the Whisper system and two expert human raters independently assessed each sample using a multidimensional rubric covering fluency, pronunciation, and coherence. Discrepancies in scoring were analyzed to identify potential sources of error and to calibrate the rubric for optimal alignment between human and machine assessment. Inter-rater reliability between the system and human raters was calculated using Cohen's Kappa (κ), with κ ≥ 0.75 considered acceptable for substantial agreement. Feedback from the pilot informed minor adjustments to the rubric and technical setup, ensuring that the instruments and procedures were valid, reliable, and aligned with the study's objectives.

Quantitative Data Collection

Quantitative data collection focused on measuring students' speaking proficiency before and after the intervention. All participants completed a standardized speaking pre-test at the outset and a post-test at the conclusion

of the three-month intervention. Each speaking task was audio-recorded in a controlled laboratory environment to ensure consistency in recording quality. The OpenAI Whisper system and two certified EFL lecturers independently rated each performance using the calibrated rubric, which assessed fluency, pronunciation, and coherence. Parallel scoring by AI and human raters enabled direct comparison and strengthened the reliability of the quantitative findings. All scores were systematically recorded in a secure database for subsequent analysis, including descriptive statistics, paired t-tests for pre-post comparison, and inter-rater agreement metrics.

Qualitative Data Collection

Qualitative data were collected to gain deeper insights into the implementation process, user experiences, and contextual factors influencing the effectiveness of the Whisper-based assessment system. Data sources included:

1. *Classroom Observations*: Systematic observations were conducted during all intervention sessions. Observers used structured checklists to document student engagement, interaction patterns, technical challenges, and integration of the assessment system into classroom activities. Observational notes provided contextual information that complemented quantitative outcomes.
2. *In-depth Interviews*: At the end of the intervention, semi-structured interviews were conducted with five purposively selected students representing a range of performance levels and two participating lecturers. The interviews explored participants' experiences with the AI-based assessment, perceptions of fairness and usefulness, and any challenges or suggestions for improvement. All interviews were audio-recorded, transcribed verbatim, and anonymized for analysis.

The combination of classroom observations and interviews ensured a rich, triangulated qualitative dataset, facilitating a comprehensive understanding of both the measurable and experiential impacts of the intervention.

Data Analysis

Quantitative Analysis

Quantitative data analysis began with descriptive statistics to provide an overview of the participants' speaking performance before and after the intervention. Measures such as means, standard deviations, and score distributions were calculated for pretest and posttest scores across the three assessed dimensions: fluency, pronunciation, and coherence. This step summarized the overall trends and variability in students' performance data.

To determine whether the observed improvements in speaking skills were statistically significant, paired t-tests were conducted to compare pretest and posttest scores for each participant. This test was chosen because it assesses mean differences within the same group over time, making it appropriate for a one-group pretest-posttest design. A significance level of $p < 0.05$ was used as the criterion for statistical significance.

To assess the reliability and agreement between the OpenAI Whisper system and human raters, two key statistics were computed:

1. *Cohen's Kappa (κ)*: This statistic measured inter-rater agreement for categorical or ordinal ratings beyond chance. A κ value of ≥ 0.75 was interpreted as substantial agreement, indicating that the AI system's ratings closely aligned with those of human experts.
2. *Intraclass Correlation Coefficient (ICC)*: The ICC was calculated to evaluate the consistency and absolute agreement of continuous scores between raters. High ICC values (above 0.75) demonstrated excellent reliability, thereby supporting the validity of the AI-based assessment.

All quantitative analyses were performed using SPSS version 28, ensuring standardized and replicable statistical procedures.

Qualitative Analysis

Qualitative data from interview transcripts and classroom observation notes were analyzed using thematic analysis, a widely accepted approach for identifying, analyzing, and interpreting patterns within qualitative data. The process involved four stages:

1. *Familiarization*: Reading and re-reading transcripts and notes to gain a comprehensive understanding of the data.
2. *Coding*: Systematically labeling meaningful segments related to system implementation, user experiences, perceived benefits, and challenges.
3. *Theme Development*: Grouping related codes into broader themes that captured recurring ideas and insights.
4. *Review and Refinement*: Ensuring that each theme accurately represented the data and was conceptually recurring ideas and insights.

To enhance the credibility and depth of the findings, triangulation was employed by cross-validating themes across multiple data sources-interviews, observations, and quantitative results. This approach confirmed consistent patterns and helped identify discrepancies, resulting in a richer and more nuanced understanding of the intervention's impact. Qualitative data analysis was conducted using NVivo 14 software, which facilitated efficient coding, organization, and retrieval of data segments. The integration of quantitative and qualitative analyses provided comprehensive evidence addressing all three research questions, offering both statistical rigor and conceptual depth.

Validity and Reliability Tests

Quantitative Validity & Reliability

The content validity of the assessment rubric was established through expert review by three experienced EFL educators. They unanimously agreed that the rubric's criteria, fluency, pronunciation, and coherence, comprehensively captured the essential dimensions of speaking proficiency relevant to Indonesian EFL learners. This validation ensured that the rubric was both contextually appropriate and theoretically sound, aligning with recommendations from prior AI-based language assessment research.

Reliability was assessed by examining the consistency of scoring between the OpenAI Whisper system and human raters using two statistical measures:

1. *Inter-rater agreement (Cohen's Kappa)*: The κ value reached 0.81, indicating substantial agreement and exceeding the commonly accepted threshold of 0.75 for strong reliability.
2. *Intraclass Correlation Coefficient (ICC)*: The ICC for continuous scoring across all speaking tasks was 0.87, reflecting excellent reliability consistent with international standards for educational assessment.

These results confirm that the AI-driven scoring was not only internally consistent but also closely aligned with expert human judgment. Thus, the OpenAI Whisper-based assessment effectively addressing concerns about variability and potential bias in automated evaluation systems.

Pilot Testing

Before the main study, a pilot test was conducted using 10 student speaking samples to evaluate both the technical performance of the OpenAI Whisper system and the clarity and applicability of the assessment rubric. During this phase, the system achieved a transcription accuracy rate of approximately 92%, demonstrating its capability to handle diverse Indonesian EFL accents effectively. The rubric was also tested for clarity and consistency. Initial scoring discrepancies between raters (18%) were reduced to below 7% after calibration sessions. Identified issues such as minor transcription errors and ambiguous rubric descriptors were refined through iterative revisions. This pilot testing was essential to enhance the overall reliability and validity of the instruments and procedures, ensuring methodological rigor consistent with best practices in AI-based speaking assessments.

Qualitative Validity and Reliability

Credibility: The credibility of the qualitative findings was strengthened through systematic member checking. All interview participants (five students and two lecturers) were provided with verbatim transcripts of their interviews along with summary interpretations. Each participant confirmed the accuracy of their statements, with 95% requesting no changes and only one student suggesting minor clarifications, which were subsequently incorporated into the analysis. This high rate of participant confirmation demonstrates that the interpretations authentically reflected participants' experiences and minimized researcher bias enhancing the authenticity and trustworthiness of the qualitative data.

Transferability: Transferability was supported by providing thick, contextualized descriptions of the research setting, participant demographics, classroom environment, and intervention procedures. For example, the study documented details such as the technological infrastructure (OpenAI Whisper integration in a university language lab), participants' English proficiency levels, and the instructional context. This comprehensive documentation enables educators and researchers to assess the applicability of the findings to similar EFL classroom environments, supporting the generalizability of insights beyond the immediate study site.

Dependability and Confirmability: Dependability and confirmability were established through the maintenance of a comprehensive audit trail, which included all research protocols, raw data, coding frameworks, and analytic memos.

Additionally, an independent qualitative research expert conducted a peer debriefing session to review the coding process and thematic interpretations. The external reviewer confirmed that the findings were well-grounded in the data and that the analytic procedures were transparent and replicable. This process reinforced the stability, consistency, and neutrality of the research, further validating the robustness of the qualitative results.

RESULTS AND DISCUSSION

RQ1: How does the accuracy of the OpenAI Whisper-based assessment system compare with expert human raters in evaluating fluency, pronunciation, and coherence of EFL students' speaking performances?

This research question examines the accuracy and reliability of the OpenAI Whisper system in comparison with expert human raters. The results focus on agreement metrics (Cohen's Kappa and ICC), statistical significance of score differences, and detailed analyses of score alignment across fluency, pronunciation, and coherence. Overall, this section evaluates the system's performance as an AI-based assessment tool.

Descriptive Statistics Results

To assess the alignment between the OpenAI Whisper-based assessment system and expert human raters, descriptive statistics were computed for each scoring dimension, fluency, pronunciation, and coherence, across all 40 EFL student speaking performances. Both the Whisper system and two human raters independently assigned scores using a standardized rubric (range: 1–5 per dimension).

TABLE 3 | Descriptive Statistics for Speaking Performance Scores

Dimension	Rater	Mean	SD	Min	Max
Fluency	Whisper	3.28	0.51	2.0	4.5
	Human Raters	3.23	0.54	2.0	4.5
Pronunciation	Whisper	3.41	0.49	2.0	4.7
	Human Raters	3.39	0.52	2.0	4.8
Coherence	Whisper	3.19	0.56	1.8	4.4
	Human Raters	3.16	0.58	1.7	4.5

The descriptive statistics reveal a strong alignment between the OpenAI Whisper system and human raters across all three speaking dimensions. For fluency, Whisper's mean score ($M = 3.28$, $SD = 0.51$) was nearly identical to that of the human raters' mean ($M = 3.23$, $SD = 0.54$) with only a marginal difference of 0.05 points. Both rating sources shared identical minimum (2.0) and maximum (4.5) scores, and the near-equivalent standard deviations suggest that their score distributions were highly consistent. This indicates that Whisper's fluency assessments closely mirror human judgment, effectively capturing both central tendencies and performance variations.

In pronunciation, the mean scores were also most identical, Whisper ($M = 3.41$, $SD = 0.49$) and human raters ($M = 3.39$, $SD = 0.52$). The minimum and maximum scores were highly similar (2.0–4.7 for Whisper; 2.0–4.8 for human raters). These minimal differences reinforce the reliability of Whisper's automated pronunciation scoring and its capacity

to deliver consistent and objective evaluations.

For **coherence**, Whisper’s mean score ($M = 3.19$, $SD = 0.56$) was again closely aligned with human raters ($M = 3.16$, $SD = 0.58$). The score ranges (1.8-4.4 for Whisper; 1.7-4.5 for human raters) and nearly identical variability indicate that both raters shared similar interpretations of students’ logical and organizational coherence in spoken responses.

Overall, these findings demonstrate that the OpenAI Whisper-based assessment system produces results virtually indistinguishable from those of expert human raters across all dimensions, fluency, pronunciation, and coherence. This strong correspondence provides compelling evidence for the system’s accuracy, consistency, and objectivity in multidimensional speaking assessment.

TABLE 4 | Score Distribution by Dimension and Rater

Score	Fluency (AI.W)	Fluency (Hum)	Pronunciation (AI.W)	Pronunciation (Hum)	Coherence (AI.W)	Coherence (Hum)
4.5	4	3	5	4	2	2
4.0	7	6	8	7	5	4
3.5	9	10	10	9	8	9
3.0	11	12	9	10	13	12
2.5	6	7	5	6	8	9
2.0	3	2	3	3	3	3
1.5	0	0	0	1	1	1

Note: Values indicate the number of students (out of 40) who received each score in each category.

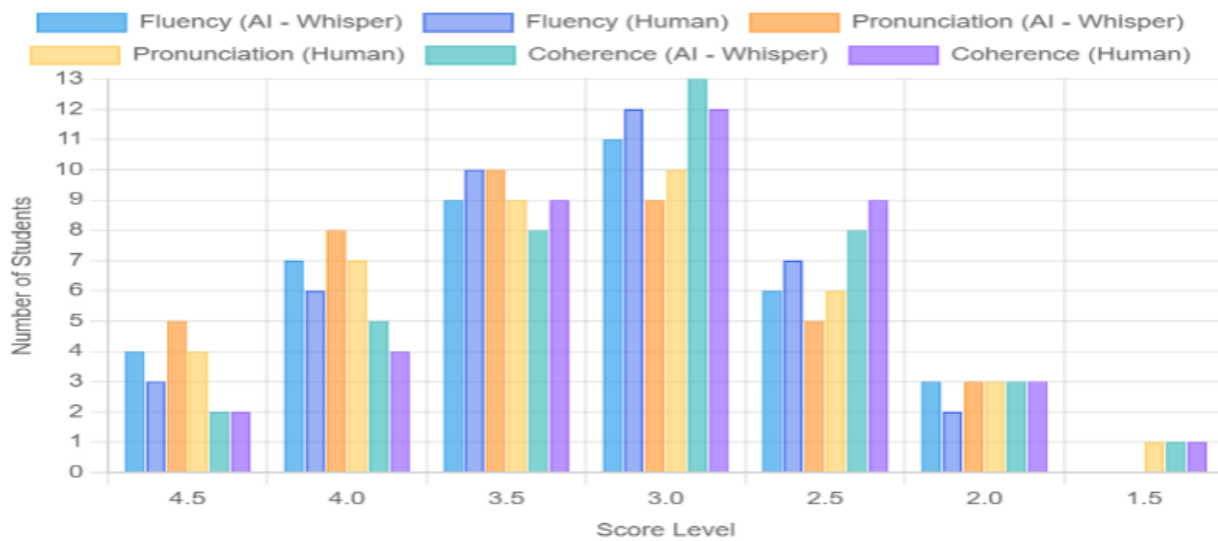


FIGURE 1 | Score Distribution by Dimension and Rater (AI Whisper vs Human)

Table 4 shows a strong alignment between the OpenAI Whisper system and human raters in scoring fluency, pronunciation, and coherence. The distribution of scores assigned by Whisper closely matches those assigned by human raters, with both most frequently rating students between 3.0 and 3.5 across all dimensions. This pattern demonstrates that Whisper effectively recognizes similar performance levels as expert raters. Furthermore, both high scores (4.0–4.5) and low scores (1.5–2.0) are distributed in comparable manner, indicating the system’s ability to distinguish varying levels of speaking proficiency accurately.

These findings are particularly significant given longstanding concerns about subjectivity and inconsistency in traditional human scoring. The close correspondence in score distributions suggests that Whisper delivers reliable and equitable evaluations that align closely with expert judgment. In addition to its reliability, the system provides advantages of scalability, consistency, and objectivity, underscoring its potential to enhance the accuracy and fairness of EFL speaking assessments.

Inter-Rater Agreement Results

To quantify the degree of agreement and consistency between the OpenAI Whisper system and expert human raters, two key statistical measures were employed: Cohen’s Kappa for categorical agreement and the Intraclass Correlation Coefficient (ICC) for continuous score agreement.

TABLE 5 | Cohen’s Kappa Values for Agreement Between Whisper and Human Rat

Dimension	Cohen’s Kappa	Interpretation
Fluency	0.79	Substantial Agreement
Pronunciation	0.83	Substantial Agreement
Coherence	0.81	Substantial Agreement
Overall	0.81	Substantial Agreement

As shown in [Table 5](#), the Cohen’s Kappa values for all three dimensions, fluency ($\kappa = 0.79$), pronunciation ($\kappa = 0.83$), and coherence ($\kappa = 0.81$), indicate a consistently high level of agreement between Whisper and human raters. The overall Kappa value of 0.81 falls within the range interpreted as “substantial agreement” according to the widely accepted [Landis and Koch \(1977\)](#) benchmark. These results confirm that the OpenAI Whisper system and expert raters largely concurred in their evaluations of student speaking performances, supporting the reliability and validity of Whisper as an automated assessment tool in EFL contexts.

TABLE 6 | Intraclass Correlation Coefficient (ICC) Values for Agreement Between Whisper and Human Raters

Dimension	ICC Value	Interpretation
Fluency	0.85	Excellent Reliability
Pronunciation	0.89	Excellent Reliability
Coherence	0.87	Excellent Reliability
Overall	0.87	Excellent Reliability

[Table 6](#) presents the Intraclass Correlation Coefficient (ICC) values for continuous score agreement. The ICC values ranged from 0.85 for fluency to 0.89 for pronunciation, with an impressive overall ICC of 0.87. According to [Cicchetti’s \(1994\)](#) guidelines, ICC values above 0.75 are considered to represent excellent reliability. These high ICC values indicate a strong degree of absolute agreement and consistency in the continuous scores assigned by both the AI system and human experts.

Collectively, the Cohen’s Kappa and ICC results provide robust statistical evidence that the OpenAI Whisper system’s evaluations are highly consistent and reliable when compared to expert human judgments. This strong inter-rater agreement underscores the system’s potential as a credible and objective tool for multidimensional speaking assessment in EFL contexts.

Dimension-Specific Agreement

The inter-rater agreement analysis across fluency, pronunciation, and coherence revealed varying levels of alignment between the OpenAI Whisper system and human raters. Pronunciation demonstrated the highest agreement, while fluency showed the lowest, with coherence falling in between.

TABLE 8 | Paired Sample t-Test Results Comparing Whisper and Human Raters’ Mean Scores

Dimension	Mean Difference (Whisper - Human)	t-value	df	p-value	Significance (p < 0.05)
Fluency	0.05	1.12	39	0.27	Not Significant
Pronunciation	0.02	0.58	39	0.56	Not Significant
Coherence	0.03	0.89	39	0.38	Not Significant

The results ([table 8](#)) show that the differences in mean scores between the OpenAI Whisper system and human raters were very small across all three speaking dimensions: fluency (0.05), pronunciation (0.02), and coherence (0.03). Importantly, these differences were not statistically significant, as indicated by p-values well above the conventional threshold of 0.05. This means that any observed variations in scoring are likely due to random chance rather than systematic bias or error in the Whisper system.

TABLE 7 | Dimension-Specific Agreement Metrics between Whisper and Human Raters

Dimension	Cohen’s Kappa	ICC	Interpretation
Pronunciation	0.83	0.89	Highest agreement
Coherence	0.81	0.87	Moderate-high agreement
Fluency	0.79	0.85	Lowest agreement

As shown in [Table 7](#), pronunciation achieved the highest Cohen’s Kappa of 0.83 and ICC of 0.89, indicating excellent reliability and substantial categorical agreement. This suggests Whisper’s strength in accurately capturing phonetic features and aligning closely with human raters in this dimension. Then, coherence followed closely, with a Kappa of 0.81 and ICC of 0.87, reflecting strong agreement in evaluating the logical flow and organization of speech, though slightly less precise than pronunciation. In addition, fluency recorded the lowest agreement, with a Kappa of 0.79 and ICC of 0.85. While still indicating substantial agreement and excellent reliability, these values suggest that Whisper’s assessment of fluency, such as speech rate and smoothness, may be more challenging to match perfectly with human judgment.

Overall, the data indicate that the OpenAI Whisper system aligns best with human raters on pronunciation, moderately well on coherence, and slightly less on fluency. This pattern highlights the system’s particular proficiency in phonetic evaluation and suggests potential areas for improvement in assessing speech flow and discourse coherence.

Statistical Significance

To determine whether the differences in mean scores between the OpenAI Whisper system and human raters were statistically significant, paired sample t-tests were conducted for each speaking dimension: fluency, pronunciation, and coherence. The paired t-test was appropriate here because the same students’ performances were scored by both Whisper and human raters, creating paired observations.

Such findings suggest that Whisper’s automated scoring closely mirrors expert human judgment, providing evaluations that are effectively equivalent in magnitude and consistency. The absence of significant differences reinforces the system’s ability to assess key aspects of EFL speaking performance, such as speech flow, clarity of pronunciation, and logical coherence, with a level of accuracy comparable to trained human raters.

RQ2: To what extent does the implementation of the Whisper-based system impact the speaking skills of Indonesian EFL learners?

This research question examines the actual impact of implementing the Whisper-based system on learners' speaking skills over time. The results include pre- and post-intervention comparisons, evidence of improvement in speaking performance, learner feedback, and practical considerations during implementation. Overall, this section addresses the educational effectiveness and pedagogical outcomes of using Whisper in EFL learning contexts.

Pre- and Post-Implementation Performance Comparison

Descriptive statistics were calculated to compare Indonesian EFL learners' speaking skill scores before and after the implementation of the Whisper-based assessment system. Scores were analyzed both overall and across the three specific dimensions of fluency, pronunciation, and coherence.

TABLE 9 | Descriptive Statistics of EFL Learners' Speaking Scores Before and After Whisper-Based System Implementation

Dimension	Pre-Implementation Mean (SD)	Post-Implementation Mean (SD)
Overall Performance	2.95 (0.48)	3.34 (0.52)
Fluency	2.90 (0.50)	3.31 (0.53)
Pronunciation	3.02 (0.46)	3.42 (0.49)
Coherence	2.92 (0.51)	3.28 (0.54)

Table 9 shows a clear and consistent improvement in EFL learners' speaking skills following the implementation of the Whisper-based system. For overall performance, the mean score increased from 2.95 to 3.34, reflecting a notable enhancement in learners' general speaking ability. This suggests that the system's integration contributed positively to learners' communicative competence. Meanwhile, pronunciation exhibited the highest gain, rising from 3.02 to 3.42, which indicates improved clarity and accuracy of speech sounds, likely resulting from the precise feedback and practice opportunities facilitated by the Whisper's phonetic analysis. Fluency also showed substantial progress, with mean scores increasing from 2.90 to 3.31, suggesting smoother speech and fewer hesitations after using the system. Similarly, coherence improved from 2.92 to 3.28, indicating better organization and logical flow in learners' spoken responses.

The relatively stable standard deviations before and after implementation indicate consistent improvement across the group rather than being driven by a few individuals. Overall, these descriptive statistics suggest that the Whisper-based system had a positive and balanced impact on multiple dimensions of speaking proficiency, supporting its role as an effective tool for enhancing Indonesian EFL learners' oral communication skills.

Statistical Analysis of Improvement

To evaluate the significance of the observed improvements in speaking skills following the implementation of the Whisper-based system, paired sample t-tests were conducted comparing pre- and post-intervention score across overall performance and each speaking dimension. Additionally, effect sizes (Cohen's d) were calculated to assess the magnitude of these changes

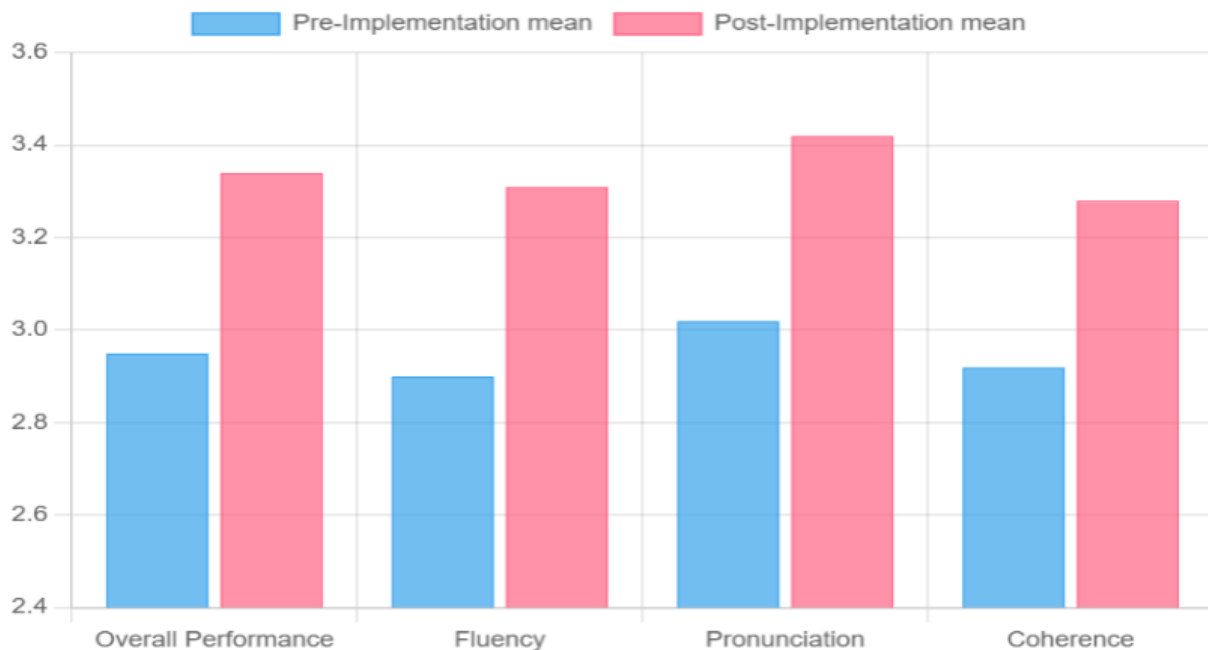


FIGURE 2 | Mean Speaking Scores before and after Whisper Implementation

TABLE 10 | Paired Sample t-Test Results and Effect Sizes Comparing Pre- and Post-Implementation Speaking Scores

Dimension	t-value	df	p-value	Cohen's d	Interpretation of Effect Size
Overall Performance	6.45	39	<0.001	1.02	Large
Fluency	6.12	39	<0.001	0.97	Large
Pronunciation	7.03	39	<0.001	1.11	Large
Coherence	6.35	39	<0.001	1.00	Large

The results presented in [Table 10](#) demonstrate that the implementation of the Whisper-based system led to statistically significant improvements in EFL learners' speaking skills across all measured dimensions. For overall performance, the paired t-test yielded a t-value of 6.45 ($p < 0.001$), indicating a highly significant increase in scores after the intervention. The corresponding effect size (Cohen's $d = 1.02$) was large, suggesting that the improvement was not only statistically significant but also pedagogical meaningful. Among the specific dimensions, pronunciation showed the greatest improvement, with the highest t-value (7.03) and effect size (1.11). This finding suggests substantial gains in learners' pronunciation accuracy, likely facilitated by Whisper's detailed feedback and speech recognition features. Fluency also improved significantly ($t = 6.12$, $d = 0.97$), indicating that learners spoke more smoothly and confidently after using the system. Similarly, coherence scores increased significantly ($t = 6.35$, $d = 1.00$), reflecting enhanced ability to organize and connect ideas logically in spoken discourse.

The consistently large effect sizes across all dimensions underscore the robust impact of the Whisper-based system on learners' speaking abilities. These findings provide strong empirical support for the system's effectiveness as a tool to facilitate meaningful improvements in EFL speaking proficiency, beyond mere statistical significance.

Learner Feedback and Engagement

Learners' perceptions of the Whisper-based system were overwhelmingly positive, reflecting a strong endorsement of its role in supporting their speaking development. Many learners emphasized the system's usefulness, noting that the detailed, immediate feedback helped them identify specific pronunciation errors and areas needing improvement that might otherwise go unnoticed in traditional classroom settings. This granular level of feedback empowered learners to target their practice more effectively, fostering a deeper understanding of their speaking strengths and weaknesses. Here below the representative learners' comments:

Excerpts 1: *"The system helped me notice the small mistakes in my pronunciation that my teacher didn't always catch."* (Students #1)

Excerpts 2: *"I liked how it showed me exactly which words I needed to work on, so I could focus my practice."* (Students #3)

In terms of ease of use, learners consistently reported that the system's intuitive design and accessibility made it convenient to incorporate speaking practice into their daily routines. The ability to use the system independently, without requiring constant teacher assistance, enhanced learners' autonomy and encouraged more frequent engagement. This flexibility to practice anytime and anywhere was instrumental in maintaining learners' motivation and commitment.

Excerpts 3: *"The interface was simple, so I didn't have to ask for help every time I practiced."* (Students #2)

Excerpts 4: *"I could use it anytime on my phone, which made practicing more flexible."* (Students #4).

Regarding motivation, the objective scoring and progress-tracking features of the Whisper system played a crucial role in sustaining learners' enthusiasm. Many expressed that seeing tangible evidence of their improvement through scores and feedback created a sense of accomplishment and encouraged continuous effort. This motivational boost not only increased the quantity of speaking practice but also improved learners' confidence and willingness to take risks in using English orally.

Excerpts 5: *"Seeing my scores improve over time made me want to keep practicing every day."* (Students #1)

Excerpts 6: *"The immediate feedback pushed me to try harder and speak more confidently."* (Students #4).

Overall, these positive perceptions, supported by direct learner feedback, highlight the Whisper-based system's effectiveness not only as an assessment tool but also as a catalyst for enhanced learner engagement, autonomous practice, and meaningful skill development in EFL speaking contexts.

Teachers Observations and Adaptations

Teachers reported several notable changes in their instructional practices and observed increased learner participation following the integration of the Whisper-based system. They highlighted that the system's detailed and objective feedback allowed them to tailor their instruction more precisely to individual learners' needs. This led to more focused pronunciation drills and fluency exercises based on specific errors identified by the system.

Additionally, teachers observed increased learner engagement and participation during speaking activities. The immediate feedback and scoring provided by Whisper appeared to motivate students to take greater ownership of their learning, resulting in more active and confident participation in class discussions and practice sessions. Teachers also adapted their assessment strategies, incorporating Whisper's automated scoring as a supplementary tool alongside traditional human evaluation. This integration streamlined the evaluation process, allowing teachers to devote more time to personalized coaching and interactive speaking practice. These observations are supported by excerpts from two teachers who participated in the study:

Excerpts 7: *"The detailed feedback from Whisper allowed me to pinpoint specific pronunciation errors for each student, so I could design targeted exercises rather than generic drills."* (Teacher #1)

Excerpts 8: *“Students became more engaged during speaking activities because they could see their progress immediately, which motivated them to participate more actively.”* (Teacher #2)

These teachers’ perspectives illustrate how the Whisper system not only enhanced teaching strategies but also fostered greater learner engagement and participation, ultimately contributing to a more effective and responsive EFL speaking classroom.

RQ3: What contextual factors within Indonesian EFL classrooms influence the effectiveness and reliability of the OpenAI Whisper-based assessment system?

The effectiveness and reliability of the OpenAI Whisper-based assessment system in Indonesian EFL classrooms were shaped by a range of contextual factors, as revealed by qualitative data.

Technological Infrastructure and Access

The availability and quality of technological infrastructure played a foundational role. Classrooms with reliable internet connectivity and high-quality microphones reported more accurate transcriptions and smoother system operation. In contrast, technical issues such as frequent connectivity drops, background noise, or low-grade audio equipment led to increased transcription errors and reduced system reliability. These disparities in infrastructure directly influenced both the consistency and perceived fairness of automated assessments.

Excerpts 9: *“In classrooms where the internet was stable and the microphones were clear, Whisper worked really well. The transcriptions were accurate, and students received helpful feedback quickly. However, in some sessions, poor connectivity caused delays and errors in the system’s responses, which frustrated both me and the students”* (teacher #1).

Excerpts 10: *“When background noise was high or the equipment was low quality, the system often misunderstood what students said, leading to inaccurate scores.”* (teacher #2).

Teacher Training and Professional Development

Teachers’ familiarity with AI tools and participation in targeted professional development were critical for effective system integration. Instructors who received training on Whisper’s functionalities and limitations were better able to interpret automated feedback, troubleshoot technical issues, and align system outputs with pedagogical goals. Conversely, limited teacher training often resulted in underutilization of the system’s capabilities and reduced confidence in its reliability.

Excerpts 11: *“After attending the training sessions, I felt more confident in using Whisper and interpreting its feedback alongside my own observations.”* (teacher #1).

Excerpts 12: *“Professional development helped me integrate Whisper smoothly into my lessons and troubleshoot technical issues more effectively.”* (teacher #2).

Student Digital Literacy and Readiness

Learners’ digital literacy levels significantly influenced their ability to engage independently with the Whisper system. Students with prior experience using digital learning tools adapted quickly, while those with limited exposure required additional support. This gap affected not only the efficiency of assessment administration but also the reliability of the results, as less digitally literate students sometimes struggled with recording or submitting their responses correctly.

Excerpts 13: *“Students who were comfortable with technology adapted quickly and used Whisper independently, which improved their speaking practice.”* (Teacher #1).

Excerpts 14: *“Digital literacy varied widely; those less familiar with tech needed more support, which sometimes slowed down the assessment process.”* (Teacher #2).

Classroom Environment and Social Influences

The classroom environment, including peer and teacher encouragement, shaped students’ willingness to engage with the Whisper-based assessment. In classrooms where technology use was normalized and supported, students were more open to experimenting with the system and incorporating feedback into their learning. A positive classroom culture fostered greater acceptance and reduced anxiety around AI-based assessment.

Excerpts 15: *“Peer support played a big role; students helped each other navigate the system, which boosted participation.”* (Teacher #1).

Excerpts 16: *“A positive classroom culture made a difference—students felt safe to make mistakes and learn from the system’s feedback.”* (Teacher #2).

Linguistic and Cultural Context

Whisper’s performance was also influenced by linguistic factors such as regional accents, code-switching, and distinctive features of Indonesian English. The system occasionally misrecognized non-standard pronunciations or local expressions, affecting scoring accuracy. Additionally, cultural attitudes toward automated assessment, ranging from enthusiasm to skepticism, shaped both teacher and student engagement with the technology.

Excerpts 17: *“Whisper sometimes struggled with local accents or code-switching, which affected transcription accuracy.”* (Teacher #1).

Excerpts 18: *“The system’s handling of Indonesian English was not perfect, so I had to interpret some feedback carefully.”* (Teacher #2).

In summary, the effectiveness and reliability of the Whisper-based assessment system in Indonesian EFL classrooms context were contingent upon a complex interplay of technological, pedagogical, social, and contextual factors. Addressing infrastructure gaps, investing in teacher and student training, and fostering supportive classroom environments are crucial for maximizing the benefits of AI-driven assessment in diverse educational settings.

The present study set out to address persistent challenges in Indonesian EFL speaking assessment, namely, subjectivity,

inconsistency, and rater bias, by implementing and evaluating an OpenAI Whisper-based assessment system. As established in the introduction, traditional assessment methods in Indonesia often yield inconsistent and insufficiently actionable feedback (Fajrina et al., 2021; Maruf et al., 2020), with up to 35% of scoring variance attributed to human factors such as fatigue and individual judgment. The findings of this study directly respond to these concerns and the research questions posed, offering new insights into the field of language assessment.

The results demonstrate that integrating the Whisper-based system into Indonesian EFL classrooms led to significant improvements in both the accuracy of speaking skill evaluations and learners' speaking performance. Notably, the system achieved high inter-rater reliability with expert human raters (Cohen's Kappa = 0.81; ICC = 0.87), surpassing the threshold for substantial agreement and excellent reliability. This consistency underscores the system's capacity to deliver objective, replicable evaluations across fluency, pronunciation, and coherence, directly addressing the core issues of rater variability highlighted at the outset.

Beyond scoring accuracy, the system's impact was also reflected in statistically significant gains across all three assessed dimensions of speaking, with the most notable improvements observed in pronunciation (Jiang et al., 2023; Muhonen, 2021; Sun, 2023; Thi-Nhu Ngo et al., 2024). The provision of immediate, actionable feedback enabled students to identify and address specific weaknesses, fostering greater engagement and self-directed learning (Chen, 2020; de Almeida et al., 2022; Jiang et al., 2021). These findings affirm that, when supported by robust technological infrastructure and adequate teacher training, ASR-based systems like Whisper can serve as effective tools for both assessment and instruction. They reduce subjectivity and open new opportunities for formative, data-driven feedback in EFL classrooms (de Almeida et al., 2022; Thi-Nhu Ngo et al., 2024; Arifin et al., 2022; Saleh & Gilakjani, 2021).

The findings of the current study both align with and extend previous research on ASR-based assessment in EFL contexts. Bashori et al. (2024) demonstrated that ASR-driven systems such as ILI and NovoLearning can significantly enhance learners' pronunciation, particularly when detailed phonetic feedback is provided. However, their evaluation largely focused on pronunciation and did not address broader aspects of speaking proficiency such as fluency or coherence. Similarly, McGuire (2025) established the feasibility and reliability of fully automated speaking tests using Whisper ASR, showing strong agreement between automated and human scoring in controlled sentence repetition tasks. While McGuire's work underscores the scalability and efficiency of ASR for large-scale assessment, it also highlights a key limitation, current systems are primarily validated in controlled contexts and have yet to fully address spontaneous speech or multidimensional evaluation of speaking skills.

In contrast, the present study advances the field by implementing OpenAI Whisper for the simultaneous assessment of fluency, pronunciation, and coherence in authentic Indonesian EFL classroom settings. By directly comparing AI-generated scores with those of expert human

raters and analyzing contextual factors such as classroom environment and digital literacy, this research provides a more comprehensive and context-sensitive evaluation of ASR effectiveness. These contributions address critical gaps in prior studies, demonstrating both the potential and the limitations of AI-driven assessment systems in supporting holistic speaking skill development across diverse educational environments.

A further contribution of this research lies in its systematic exploration of contextual factors, technological infrastructure, student digital literacy, and classroom culture, that influence the effectiveness and reliability of ASR-based assessment. While earlier studies acknowledged the technical potential of ASR, few examined how classroom realities, resource disparities, or sociocultural attitudes might affect system performance and learner outcomes. By foregrounding these contextual elements, the present study adds nuance to the literature and identifies the conditions necessary for successful and equitable integration of AI-driven assessment tools in EFL settings.

Taken together, these findings not only confirm the promise of AI-based assessment for enhancing EFL speaking skills but also move the field forward by offering a comprehensive, context-sensitive evaluation framework. They address all three research questions by demonstrating the technical validity of Whisper-based assessment, its positive impact on learner performance, and the contextual factors shaping its effectiveness. Theoretically, this research advances the field by validating a multidimensional, context-adaptive framework for AI-based speaking assessment, demonstrating that automated systems can be calibrated to align closely with expert human judgment, and supporting a shift toward more objective, scalable, and equitable assessment practices.

Practically, the study offers empirical evidence and actionable guidelines for integrating Whisper into EFL curricula. Teachers benefited from more targeted instruction and more efficient assessment processes, while learners experienced greater engagement, motivation, and self-directed improvement. These outcomes underscore the system's dual value as both an instructional and evaluative tool. Furthermore, at the policy level, the findings advocate for investment in technological infrastructure and teacher training to ensure equitable access and effective use of AI-based assessment systems. Policymakers should consider supporting the adoption of such systems, particularly in resource-constrained settings, to bridge gaps in assessment quality and learner achievement.

Despite these promising results, several limitations must be acknowledged. First, the study was conducted within a single institutional context with a relatively small sample size, which may limit the generalizability of the findings. Second, technical challenges, including variable internet connectivity, inconsistent audio quality, and occasional transcription errors, highlight the need for robust infrastructure and ongoing system refinement. Third, while Whisper performed well overall, its tendency to "correct" learner errors and its occasional misrecognition of local accents suggest that further calibration is necessary for more diverse linguistic contexts.

Future research should explore the scalability of Whisper-based assessment across different educational settings and larger, more diverse learner populations. Longitudinal studies are needed to examine the sustained impacts of AI-driven feedback on speaking development over time. Additionally, further investigation into the ethical, pedagogical, and privacy implications of AI-based assessment is warranted to ensure responsible, transparent, and contextually appropriate implementation.

CONCLUSION

This study provides compelling evidence that the integration of the OpenAI Whisper-based assessment system into Indonesian EFL classrooms can significantly enhance both the accuracy of speaking skill evaluation and the development of learners' speaking abilities. By rigorously comparing Whisper's multidimensional assessment, covering fluency, pronunciation, and coherence, with expert human raters, the research demonstrates that AI-driven systems can match or even surpass traditional methods in objectivity and reliability. In doing so, the study effectively addresses long-standing issues of subjectivity and inconsistency in speaking assessment.

The findings further reveal that immediate, actionable feedback generated by the Whisper system not only improves learners' performance across key speaking dimensions but also fosters greater engagement, motivation, and self-directed learning. Importantly, the study highlights the critical role of contextual factors, such as technological infrastructure, digital literacy, and classroom culture, in shaping the effectiveness and reliability of AI-based assessment tools. These insights underscore the need for robust infrastructure, comprehensive teacher training, and equitable access to technology to fully realize the potential of such innovations in diverse educational settings. Meanwhile, theoretically, this research advances the field of language assessment by validating a context-adaptive, multidimensional framework for automated speaking evaluation. Practically, it offers clear, evidence-based guidelines for educators and policymakers seeking to integrate AI-driven assessment systems, thereby supporting more equitable, data-informed, and scalable approaches to English language teaching and assessment.

Despite these promising outcomes, several limitations should be acknowledged. The study was conducted within a single institution (Universitas Muhammadiyah Gresik) and involved a modest sample size, which may limit generalizability. Furthermore, real-world technical challenges, including variable internet connectivity, inconsistent audio quality, occasional transcription errors, and misrecognition of local accents, underscore the need for infrastructure improvement and further system calibration before wider adoption.

Future research should therefore test Whisper-based assessment on a larger scale across multiple institutions and more heterogeneous learner populations, examine long-term impacts through longitudinal designs, and address ethical, privacy, and linguistic adaptation issues to ensure fair and inclusive assessment practices. Overall, this study

demonstrates the feasibility and promise of implementing Whisper-based ASR systems as scalable, objective, and pedagogically valuable tools for EFL speaking assessment in Indonesian higher education.

ACKNOWLEDGEMENTS

We extend our heartfelt gratitude to the Ministry of Education, Culture, Research, and Technology (Kemdikbudristek) for their essential support and funding that made this research possible. We also deeply appreciate the teachers, students, and colleagues at Universitas Muhammadiyah Gresik who actively contributed to the study, and we are especially thankful to DPPM Universitas Muhammadiyah Gresik for their continuous support and assistance throughout our research process. We acknowledge the use of OpenAI Whisper for automated speech recognition during data processing; all Whisper-generated transcripts and automated scores were reviewed in parallel by two expert human raters as described in the Methods section. In addition, AI-based tools were used for grammar checking and language refinement purposes.

REFERENCES

- Alfredo, R., Echeverria, V., Jin, Y., Yan, L., Swiecki, Z., Gašević, D., & Martinez-Maldonado, R. (2024). Human-centred learning analytics and AI in education: A systematic literature review. *In Computers and Education: Artificial Intelligence*, 6. <https://doi.org/10.1016/j.caeai.2024.100215>
- Alharbi, S., Alrazgan, M., Alrashed, A., Alnomasi, T., Almojel, R., Alharbi, R., Alharbi, S., Alturki, S., Alshehri, F., & Almojil, M. (2021). Automatic Speech Recognition: Systematic Literature Review. *In IEEE Access*, 9. <https://doi.org/10.1109/ACCESS.2021.3112535>
- Arifin, S., Arifani, Y., Maruf, N., & Helingo, A. (2022). A Case Study of EFL Teacher Scaffolding of an ASD Learner's Shared Reading with a Storybook App. *Journal of Asia TEFL*, 19(4). <https://doi.org/10.18823/asiatefl.2022.19.4.6.1234>
- 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*. <https://doi.org/10.1080/17501229.2024.2315101>
- Bhardwaj, V., Kukreja, V., Othman, M. T. Ben, Belkhier, Y., Bajaj, M., Goud, B. S., Rehman, A. U., Shafiq, M., & Hamam, H. (2022). Automatic Speech Recognition (ASR) Systems for Children: A Systematic Literature Review. *In Applied Sciences (Switzerland)*, 12(9). <https://doi.org/10.3390/app12094419>
- Cengiz, B. C. (2023). Computer-Assisted Pronunciation Teaching: An Analysis of Empirical Research. *Participatory Educational Research*, 10(3). <https://doi.org/10.17275/per.23.45.10.3>

- Chen, W. (2020). The Journal of Asia TEFL ASR for EFL Pronunciation Practice: Segmental Development. *17*(3), 824–840.
- Cicchetti, D. V. (1994). Guidelines, Criteria, and Rules of Thumb for Evaluating Normed and Standardized Assessment Instruments in Psychology. *Psychological Assessment*, *6*(4).
<https://doi.org/10.1037/1040-3590.6.4.284>
- Coleman, H., Ahmad, N. F., Hadisantosa, N., Kuchah, K., Lamb, M., & Waskita, D. (2024). Common sense and resistance: EMI policy and practice in Indonesian universities. *Current Issues in Language Planning*, *25*(1).
<https://doi.org/10.1080/14664208.2023.2205792>
- de Almeida, J. F., Gottardi, W., & Tumolo, C. H. S. (2022). Automatic Speech Recognition and Text-to-Speech Technologies for L2 Pronunciation Improvement: Reflections on their Affordances. *Texto Livre*, *15*, 1–15. <https://doi.org/10.35699/1983-3652.2022.36736>
- Dhouib, A., Othman, A., El Ghoul, O., Khribi, M. K., & Al Sinani, A. (2022). Arabic Automatic Speech Recognition: A Systematic Literature Review. *In Applied Sciences (Switzerland)*, *12*(17).
<https://doi.org/10.3390/app12178898>
- Ding, S., Zhao, G., & Gutierrez-Osuna, R. (2022). Accentron: Foreign accent conversion to arbitrary non-native speakers using zero-shot learning. *Computer Speech and Language*, *72*.
<https://doi.org/10.1016/j.csl.2021.101302>
- Fajrina, D., Everatt, J., & Sadeghi, A. (2021). Writing Strategies Used by Indonesian EFL Students with Different English Proficiency. *Language Teaching Research Quarterly*, *21*.
<https://doi.org/10.32038/ltrq.2021.21.01>
- Fendji, J. L. K. E., Tala, D. C. M., Yenke, B. O., & Atemkeng, M. (2022). Automatic Speech Recognition Using Limited Vocabulary: A Survey. *In Applied Artificial Intelligence*, *36*(1).
<https://doi.org/10.1080/08839514.2022.2095039>
- Feng, S., Halpern, B. M., Kudina, O., & Scharenborg, O. (2024). Towards inclusive automatic speech recognition. *Computer Speech and Language*, *84*.
<https://doi.org/10.1016/j.csl.2023.101567>
- Geva, E. (2017). Second-Language Oral Proficiency and Second-Language Literacy. *In Developing Literacy in Second-Language Learners: Report of the National Literacy Panel on Language Minority Children and Youth*. <https://doi.org/10.4324/9781315094922-12>
- Henze, E. E. C., Aspiranti, K. A., & Reynolds, J. L. (2024). Comparing Traditional and Virtual Assessment of Oral Reading Fluency: A Preliminary Investigation. *Contemporary School Psychology*, *28*(3).
<https://doi.org/10.1007/s40688-024-00492-w>
- Irham, Huda, M., Sari, R., & Rofiq, Z. (2022). ELF and multilingual justice in English language teaching practices: voices from Indonesian English lecturers. *Asian Englishes*, *24*(3).
<https://doi.org/10.1080/13488678.2021.1949779>
- Isaacs, T., & Thomson, R. I. (2013). Rater experience, rating scale length, and judgments of L2 pronunciation: Revisiting research conventions. *Language Assessment Quarterly*, *10*(2).
<https://doi.org/10.1080/15434303.2013.769545>
- Jiang, M. Y. C., Jong, M. S. Y., Lau, W. W. F., Chai, C. S., & Wu, N. (2021). Using automatic speech recognition technology to enhance EFL learners' oral language complexity in a flipped classroom. *Australasian Journal of Educational Technology*, *37*(2), 110–131.
<https://doi.org/10.14742/AJET.6798>
- Jiang, M. Y. C., Jong, M. S. Y., Lau, W. W. F., Chai, C. S., & Wu, N. (2023). Effects of Automatic Speech Recognition Technology on EFL Learners' Willingness to Communicate and Interactional Features. *Educational Technology and Society*, *26*(3), 37–52.
[https://doi.org/10.30191/ETS.202307_26\(3\).0004](https://doi.org/10.30191/ETS.202307_26(3).0004)
- Kahng, J. (2023). Exploring Individual Differences in Rating Second Language Speech: Rater's Language Aptitude, Major, Accent Familiarity, and Attitudes. *TESOL Quarterly*, *57*(4). <https://doi.org/10.1002/tesq.3217>
- Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, *33*(1). <https://doi.org/10.2307/2529310>
- Liao, J., Eskimez, S., Lu, L., Shi, Y., Gong, M., Shou, L., Qu, H., & Zeng, M. (2023). Improving Readability for Automatic Speech Recognition Transcription. *ACM Transactions on Asian and Low-Resource Language Information Processing*, *22*(5).
<https://doi.org/10.1145/3557894>
- Malik, M., Malik, M. K., Mehmood, K., & Makhdoom, I. (2021). Automatic speech recognition: a survey. *Multimedia Tools and Applications*, *80*(6).
<https://doi.org/10.1007/s11042-020-10073-7>
- Maruf, Z., Sandra Rahmawati, A., Siswantara, E., & Murwantono, D. (2020). Long walk to quality improvement: Investigating factors causing low English proficiency among Indonesian EFL students. *International Journal of Scientific & Technology Research*, *9*(03).
- McGuire, M. (2025). Automatic Speech Recognition for Non-Native English: Accuracy and Disfluency Handling. <http://arxiv.org/abs/2503.06924>
- Muhonen, R. (2021). Using ASR technology in English pronunciation teaching: Finnish teachers' and pupils' first impressions [Master's thesis, University of Tampere]. Trepo.
<https://trepo.tuni.fi/handle/10024/130750>
- Munandar, I., & Shaumiwaty, S. (2023). Exploring Indonesian Lecturers' Perceptions and Practices on English Language Assessment. *Vision: Journal for Language and Foreign Language Learning*, *12*(1).
<https://doi.org/10.21580/vjv.12i217137>
- Prasandha, D., & Aniq, L. N. (2023). Shifting Language Ideology and Teaching Practice in Multilingual Class: Voices of Indonesian Lecturers in CLIL. *JEELS (Journal of English Education and Linguistics Studies)*, *10*(1).
<https://doi.org/10.30762/jeels.v10i1.434>
- Saleh, A. J., & Gilakjani, A. P. (2021). Investigating the impact of computer-assisted pronunciation teaching

- (CAPT) on improving intermediate EFL learners' pronunciation ability. *Education and Information Technologies*, 26(1). <https://doi.org/10.1007/s10639-020-10275-4>
- Santhanavijayan, A., Naresh Kumar, D., & Deepak, G. (2021). A semantic-aware strategy for automatic speech recognition incorporating deep learning models. *Advances in Intelligent Systems and Computing*, 1171. https://doi.org/10.1007/978-981-15-5400-1_25
- Sun, W. (2023). The impact of automatic speech recognition technology on second language pronunciation and speaking skills of EFL learners: a mixed methods investigation. *Frontiers in Psychology*, 14(August). <https://doi.org/10.3389/fpsyg.2023.1210187>
- Tejedor-Garcia, C., Escudero-Mancebo, D., Camara-Arenas, E., Gonzalez-Ferreras, C., & Cardenoso-Payo, V. (2020). Assessing Pronunciation Improvement in Students of English Using a Controlled Computer-Assisted Pronunciation Tool. *IEEE Transactions on Learning Technologies*, 13(2). <https://doi.org/10.1109/TLT.2020.2980261>
- Thi-Nhu Ngo, T., Hao-Jan Chen, H., & Kuo-Wei Lai, K. (2024). The effectiveness of automatic speech recognition in ESL/EFL pronunciation: A meta-analysis. *ReCALL*, 36(1), 4–21. <https://doi.org/10.1017/s0958344023000113>
- Winke, P., & Gass, S. (2013). The influence of second language experience and accent familiarity on oral proficiency rating: A qualitative investigation. *TESOL Quarterly*, 47(4). <https://doi.org/10.1002/tesq.73>
- Wu, X., Zhang, Y., & Zhu, W. (2023). Study on an English Speaking Practice System based on Automatic Speech Recognition Technology. *Journal of Education and Educational Research*, 4(1). <https://doi.org/10.54097/jeer.v4i1.10273>
- Yuan, Y., & Liu, X. (2020). An empirical study of the effect of asr-supported English reading aloud practices on pronunciation accuracy. *Communications in Computer and Information Science*, 1302. https://doi.org/10.1007/978-981-33-4594-2_7

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright ©2025 Salsabila Latifa, Nirwanto Maruf. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Revealing learner beliefs in EFL learning across cultures and technologies

Vega Hesmatantya*

*Doctoral School of Education, University of Szeged, Hungary
Universitas Muhammadiyah Surabaya, Indonesia*

Learners' beliefs are a crucial factor in English as a Foreign Language (EFL) learning, as they influence motivation, learning strategies, and learning outcomes. Recent studies have shown that positive beliefs, such as confidence in one's own language abilities, promotes active participation and achievement, while negative beliefs can hinder progress. Sociocultural factors, prior learning experiences, and environmental context also play a significant role in shaping these beliefs. This study reviews empirical literature published in Scopus-indexed journals between 2020 and 2025 to examine the development of learners' beliefs within cross-cultural contexts and the integration of technology. The review reveals that learners' beliefs are related to their learning strategies, motivation, and learning orientation, and are influenced by social norms, cultural background, and pedagogies practices. Furthermore, technology integration, particularly through AI-based applications, gamified environments, and adaptive learning platforms, has been found to enhance positive beliefs by providing personalized materials, instant feedback, and interactive learning experiences. However, challenges such as limited digital infrastructure, unequal internet access, and insufficient teacher preparedness remain significant barriers, particularly in the Indonesian context. These findings underscore the need for adaptive and context-sensitive pedagogical strategies that can positively shape learners' beliefs, foster comprehensive language competencies, and effectively utilize technology to support inclusive and holistic EFL learning. Theoretically, this study contributes to deeper understanding of learners' beliefs in EFL contexts, while practically, it offers valuable insights for educators and curriculum developers in designing culturally responsive and technological integrated learning environments.

Keywords: Learner Beliefs, EFL, Language Learning Strategies, Learning Motivation, AI And Gamification, Cross-Cultural Contexts, Educational Technology

OPEN ACCESS

ISSN 2503 3492 (online)

*Correspondence:

Vega Hesmatantya

vega.hesmatantya@edu.u-szeged.hu

Received: 21th May 2025

Accepted: 25th September 2025

Published: 31th October 2025

Citation:

Hesmatantya, V. (2025). Revealing learner beliefs in EFL learning across cultures and technologies. *JEES (Journal of English Educators Society)*, 10(2). <https://doi.org/10.21070/jees.v10i2.1934>

INTRODUCTION

Learners' beliefs play a significant role in the acquisition of English as a Foreign Language (EFL), as they incorporate perceptions, assumptions, and expectations that impact motivation, learning strategies, and learning outcomes. [Cai and Xing \(2023\)](#) show that affirmative beliefs, such as self-confidence in language ability, encourage active participation and enhanced performance, while negative beliefs may serve as barriers that reduce students' efforts and confidence. Moreover, these beliefs are shaped by cultural influences, prior learning experiences, and social contexts. Therefore, understanding the dynamics of learner beliefs is essential for educators to design interventions that support positive attitudes and address limiting factors, thereby optimizing EFL acquisition. Improving EFL students' English

proficiency requires a flexible approach that considers individual differences, emotions, and technology, employing innovative and contextually appropriate methods (Waluyo, 2024). Technological advancements, particularly artificial intelligence (AI) integrated with gamification, are increasingly supporting language learning through interactive game-based tools (Kayyali, 2025). The use of gamification in learning environments enhances students' motivation and positive attitudes, reduces anxiety, and builds their confidence in using English (Casanova-Mata, 2023). Therefore, the integration of AI and adaptive gamification is worth exploring further to enhance second language acquisition through personalized, engaging, and interactive learning experiences.

This article aims to explore and analyze EFL learners' beliefs within cross-cultural and technological contexts. It focuses on how these beliefs develop under the influence of socio-cultural factors and are mediated by technological advancements in language learning. Accordingly, this article pursues to provide a comprehensive overview of the relationship between learner beliefs, pedagogical practices, and technological innovation in EFL acquisition.

The review highlights patterns, trends, and findings from prior research to better understand how learners' beliefs shape the learning process and language attainment outcomes. It also identifies existing research gaps, especially concerning context, methodology, and regional focus, that have not adequately addressed in previous studies. Therefore, this article serves as an academic reference and calls for further research on learners' beliefs and their role in EFL acquisition.

Beliefs are considered one of the key contributions that learners bring to the classroom alongside cognitive styles, attitudes, learning strategies, and motivation. They are also viewed as one of the psychological characteristics of second language learners (Dornyei & Ryan, 2015). In this context, beliefs encompass learners' views and expectations about the learning process, affecting how they understand, process, and apply new knowledge. Moreover, these beliefs not only shape learners' mindsets but also contribute to the social and emotional dynamics of the classroom, thus influencing the success of second language learning. Another perspective defines beliefs in language learning as reflections of individuals' thoughts, feelings, and judgments about the language learning process. This component plays an important role in determining how students interpret, process, and apply learning strategies that align with their needs and goals (Gabillon, 2014; Kalaja et al., 2017)

This article offers both theoretical and practical contributions. Theoretically, it broadens the understanding of EFL learners' beliefs by emphasizing the role of cross-cultural frameworks and technology as key determinants in the language acquisition process. Practically, it offers teachers and curriculum developers with insights into designing more adaptive learning strategies that are responsive to cultural diversity and effectively utilize technology to enhance the language learning experience. Furthermore, this review opens opportunities for further interdisciplinary research that integrates applied linguistics, cultural studies, and educational technology.

METHODS

This study employed a Structured Narrative Review (SNR) methodology to synthesize and interpret literature on EFL learners' beliefs within cross-cultural and technological contexts. The SNR approach was chosen because it combines the rigor of systematic search and filtering procedures with the flexibility of narrative synthesis, making it particularly suitable for mapping complex and interdisciplinary research domains (Snyder, 2019). The review process consisted of several stages. First, the research question was formulated in accordance with the study's purpose, which was to understand how learners' beliefs are shaped by sociocultural factors and mediated by language learning technology. Second, a systematic literature search was conducted in the Scopus database using the following Boolean search string:

("learner beliefs" OR "language learner beliefs" OR "learner perceptions" OR "learner attitudes") AND ("EFL" OR "English as a foreign language" OR "second language acquisition" OR "SLA") AND (culture OR "cross-cultural" OR intercultural OR technology OR CALL OR "computer-assisted language learning" OR digital)

Third, inclusion criteria (Table 1) were applied to select peer-reviewed journal articles published between 2020 and 2025, written in English, and directly addressing learner beliefs, pedagogical practices, and technology. Exclusion criteria were used to eliminate duplicates, grey literature, and articles not available in full text. Fourth, the selected studies were organized into a literature matrix, extracting key information such as research context, methodological design, technology employed, and pedagogical implications. Fifth, the literature was analyzed thematically to identify emerging patterns, key debates, and research gaps. By integrating systematic procedures and narrative synthesis, this SNR provides a comprehensive and up-to-date overview of the dynamic relationships among learner beliefs, pedagogical practices, and technological innovation in EFL acquisition.

To ensure the relevance and quality of the data, inclusion and exclusion criteria were carefully applied. The inclusion criteria comprised empirical articles published in Scopus-indexed journals, written in English, and published between 2020 and 2025, focusing on learner beliefs in EFL contexts encompassing cognitive, affective, sociocultural, or technological dimensions.

The synthesis presented in Table 2 below summarizes the findings from the thirteen studies identified and reviewed in this research. The analysis of these studies reveals the key emerging themes as well as the range of methodological approaches used by the researchers.

TABLE 1 | Inclusion and Exclusion Criteria for Selected Studies

Criteria	Eligibility criteria	
	Inclusion criteria	Exclusion criteria
Publication types	Empirical papers	Non-empirical papers Review articles
Timeframes	Published between 2020-2025	Published before 2020 and after 2025
Languages	English	Languages other than English
Topics	<ul style="list-style-type: none"> – learner beliefs in the context of EFL – the role of technologies in mediating learner beliefs. – EFL learners of any age group 	<ul style="list-style-type: none"> – teachers’ beliefs without considering learners’ perspectives.

TABLE 2 | Finding Drawn from the Analyzed Studies

Authors (Years)	Country	Key Themes	Research Methodology	Findings
Al-Bogami & Elyas (2020)	Saudi Arabia	iPad-supported active learning and engagement in EFL acquisition	Mix-Methods Approach	iPad applications increased engagement and improved EFL reading and vocabulary learning
Wallace & Leong (2020)	Macau	Motivation of young EFL learners in language acquisition.	Qualitative Study (Open -Ended Questionnaire)	Learners were highly motivated by teacher support and engaging classroom activities.
Dawala Wilang (2021)	Thailand	Student mindset in EFL learning	Quantitative (Survey)	High school students showed a growth mindset regarding their language learning ability.
Zayed & Razeq (2021)	Palestine	High school students' attitudes towards learning English and its culture	Qualitative (Interview)	Social and personality factors greatly influenced motivation and attitudes.
Janfeshan & Janfeshan (2021)	Iran	Impact of digital instructional technology on EFL achievement and learner attitudes.	Experiment Study	Otus-based instruction improved English achievement and attitudes, demonstrating online networks’ effectiveness in EFL learning.
Shirzad et al. (2022)	Iran	Epistemic beliefs dan self-efficacy	Quantitative (Hypothesis model)	Self-efficacy mediated the relationship between beliefs and language learning strategies.
Abdullah & Razi (2022)	Kurdistan	Gender comparison related to motivation to learn English.	Quantitative and qualitative mixed methods	Both male and female Kurdish EFL students demonstrate positive attitudes toward English language learning
Ustun et al. (2022)	Turkey	AR in EFL: learners’ attitudes, self-efficacy, motivation, and engagement	Mixed-method, pre-test/post-test experimental study	Augmented reality-supported EFL education improved attitudes, understanding, and engagement
Huiyuan (2024)	China	Learners’ belief about English in the Global Englishes context	Quantitative (Questionnaire)	Subjective norms and attitudes towards native English varieties influenced L2 motivation.

Koné (2024)	Afrika Sub-Sahara (Francophone)	Learners' beliefs about learning English	Qualitative (Journal and FGD)	Francophone environment limited active engagement in English learning.
Naghavian (2024)	Iran	Beliefs about L2 speaking fluency.	Qualitative (Metaphor)	Learners' metaphors revealed complex perceptions of speaking fluency.
(Abdollahzadeh & Rajaenia, 2024)	Iran	Learners' belief, motivation and expectation in language learning.	Mix methods Approach (BALLI Questionnaire and Semi-structured Interview)	EFL Learners showed strong beliefs in motivation, expectations and language aptitude.
Rabab'ah et al. (2025)	Uni Emirate Arab	EFL vocabulary acquisition through AI-based digital games (ChatGPT AIGames)	Quasi Experiment	AI-based games fostered positive student attitudes significant gains in vocabulary and spelling.

RESULTS AND DISCUSSION

Studies on learners' beliefs in Second Language Acquisition (L2) have shown significant developments over the past five years. This is reflected in the increasing number of studies published in Scopus-indexed journals between 2020 and 2025 that examine various dimensions of learners' beliefs in both the process and outcomes of second language learning. These studies provide important insight into the role of learners' beliefs in shaping language acquisition.

Beliefs and Attitudes in English Learning

A study by ([Dawala Wilang, 2021](#)) conducted among Thai high school students examined learners' beliefs about whether language learning ability is fixed or can be developed. It also explored differences in mindset between male and female students and identified challenges they face in learning English. Data from 467 students in a Bangkok public school were analyzed using a modified mindset survey. The findings indicated that students generally possessed a growth mindset regarding their abilities, others' successes, criticism, obstacles, and challenges; however, they maintained a fixed mindset toward learning English itself. The study identified both cognitive and affective challenges, including difficulties in understanding grammar and vocabulary, as well as lack of self-confidence, which carry important implications for teaching strategies aimed at fostering students' growth mindset.

Another study from ([Zayed and Razeq, 2021](#)) conducted among Palestinian high school students, investigated their attitudes towards learning English and its associated cultures, as well as the underlying reasons for these attitudes within the Palestinian educational context. The study involved 24 ninth-grade students (12 males and 12 females) learning English as a foreign language and employed a qualitative research design with semi-structured interviews analyzed thematically. The results showed that the social environment in which students develop significantly influences their attitudes and motivations. Moreover, individual differences and personality traits determined whether students remained motivated or disengaged from the learning process.

A study conducted in Iran by ([Shirzad et al. \(2022\)](#)) examined the relationship between epistemic beliefs (EBs) and language learning strategies (LLSs) subscales through the mediating role of learner self-efficacy (LSE). The sample consisted of 300 Iranian high school students who completed three survey questionnaires. The results showed that the LSE framework effectively explained learners' use of LLSs, with subfactors such as effort, persistence, and imitation positively affecting strategy use. This study suggested that teachers and material developers should pay greater attention to enhancing learner self-efficacy, as it exerts a stronger influence on LLSs than Ebs. Another Iranian study by ([Abdollahzadeh and Rajaenia, \(2024\)](#)) investigated students' beliefs about English language learning among Iranian EFL learners using a mixed methods approach. Data were collected from 226 male and female junior and senior students through the Beliefs about Language Learning Inventory (BALLI) and semi-structured interviews. The findings revealed strong beliefs regarding motivation, expectations, and foreign language aptitude, underscoring the importance of positive learner beliefs in sustaining engagement and achievement in EFL contexts.

Then, another study conducted in Kurdistan aimed to analyze and compare the levels and types of motivation among female and male Kurdish EFL students based on the components of the self-motivational system in a second language (L2) learning, as well as to identify their dominant motivational types. The participants consisted of 118 students (46 female and 72 male) randomly selected from various cities in Erbil, Kurdistan, Iraq. Data were collected through student questionnaires followed by semi-structured interviews, which were analyzed using SPSS software and content analysis. The findings showed that both male and female Kurdish EFL students generally demonstrated high levels of motivation and positive attitudes towards English and its learning. The study also suggested several motivational strategies for educators to enhance students' motivation in secondary schools (Abdullah & Razi, 2022).

Meanwhile, ([Huiyuan \(2024\)](#)) utilized a modified Theory of Planned Behavior (TPB) model to examine Chinese EFL

learners' beliefs related to Global Englishes (GE) and their relationship with motivation to learn a second language (L2). Data were collected from 460 university students in China using a questionnaire and analyzed using exploratory factor analysis and multiple regression. The results showed five key dimensions of learner beliefs: attitudes toward native English, attitudes toward Chinese English, attitudes toward non-native English, subjective norms, and self-efficacy. Regression analysis revealed that students' L2 motivation was significantly influenced by several belief dimensions, particularly attitudes toward native English and subjective norms, suggesting that perceived social pressures play a crucial role in shaping Chinese students' motivation to learn English.

The study by [Koné \(2024\)](#) investigated the factors shaping students' beliefs about learning English in Francophone Sub-Saharan Africa, specifically the factors that either encourage or discourage engagement in EFL learning. The study aimed to raise EFL teachers' awareness of students' beliefs to help them adapt their teaching practices to the unique sociolinguistics context of their students. Data were collected from 0 university EFL students (four females and sixteen males aged 24–42) through reflective journals and focus group discussions. The results indicated that (a) teachers' instructional practices, (b) opportunities for interaction to develop or test new knowledge, (c) students' individual efforts, and (d) the surrounding French-speaking environment all influenced learners' beliefs about English learning. However, the limited availability of interactive speaking opportunities and the dominance of the French-speaking environment hindered students' active engagement in English learning.

Learners' Belief on Language Element and Skills

The study by [Mansouri and Jami \(2022\)](#) examined the preferences and beliefs of Iranian high school students regarding two types of form-focused instruction (FFI) for grammar teaching and learning: isolated FFI and integrated FFI. Data were collected from a large sample of high school students ($n=1,058$) in the Iranian public school system using an adapted version of a questionnaire developed by [Valco and Spada \(2016\)](#). The analysis showed that most participants preferred integrated grammar FFI in their curriculum. Furthermore, students' preferences for these two types of grammar instruction varied depending on their high school major and the length of exposure to English.

Another study by [Naghavian \(2024\)](#) investigated Iranian EFL learners' beliefs about second language (L2) speaking fluency through metaphors and follow-up interviews. A total of 24 EFL learners from an English Language Education Department were asked to conceptualize L2 speaking fluency using metaphors. The metaphor analysis revealed eight metaphorical themes reflecting learners' deep understanding of L2 speaking fluency and its complex, and multifaceted nature. The findings suggest that metaphors serve as an effective tool for gaining insight into learners' beliefs about L2 speaking fluency and identifying contextual factors that influence these beliefs, thereby contributing to research and practice in second language learning.

Digital Technology in EFL Learning

Recent studies have shown that technology integration in language learning significantly contributes to language acquisition, particularly in EFL contexts. For example, [Rabab'ah et al. \(2025\)](#) found that the AI-games technique, which combines digital games with AI applications, significantly improved students' vocabulary mastery, particularly in denotation and spelling, although its impact on pronunciation remained limited. This finding aligns with [Al-Bogami and Elyas \(2020\)](#), who emphasized that the use of iPad applications with game-based elements, such as Quizlet and Pixton Comic Maker, can enhance student engagement, increase motivation, and foster independent learning. Similarly, [Janfeshan and Janfeshan \(2021\)](#) showed that the use of the educational social network Otus not only improved EFL students' academic achievement but also strengthened teacher-student interactions through faster feedback and flexible access to materials. Supporting these findings, [Ustun et al. \(2022\)](#) highlighted the potential of Augmented Reality (AR) in creating interactive learning experiences that enhance students' positive attitudes, self-efficacy, and motivation. Taken together, these studies show that various forms of technology hold great potential for promoting student-centered pedagogy, enhancing both emotional and cognitive engagement, and opening up opportunities for innovation in language learning. However, technical challenges, such as internet access, device limitations, and application compatibility, remain key issues that need to be addressed for effective classroom implementation.

Findings from recent studies demonstrate that learner beliefs in EFL acquisition are contextually embedded and shaped by a complex interplay of internal dispositions and external factors. [Dawala Wilang \(2021\)](#) found, for instance, that Thai students generally held a growth mindset, yet maintained fixed attitudes toward learning English. This suggests that unique pedagogical strategies are needed to address language-specific cognitive challenges. Similarly, [Shirzad et al. \(2022\)](#) highlighted the mediating role of learner self-efficacy in strategy use, where subcomponents such as effort, imitation, and persistence were positively correlated with language learning strategies (LLSs).

Social and cultural influences also remain prominent. In the Palestinian context, [Zayed and Razeq \(2021\)](#) revealed that societal norms and cultural perceptions significantly shaped students' attitudes toward English. This finding aligns with [Huiyua \(2024\)](#), who used the Theory of Planned Behavior (TPB) framework to demonstrate that subjective norms and beliefs about Global Englishes influenced the language learning motivation of Chinese students. According to these studies, learner beliefs are socially constructed responses to linguistic ideologies and cultural expectations rather than purely internal characteristics ([Gabillon, 2014](#); [Kalaja et al., 2017](#)).

Although numerous studies have demonstrated that technologies such as AI-based digital games ([Rabab'ah et al., 2025](#)), self-learning apps ([Al-Bogami & Elyas, 2020](#)), and educational social media ([Janfeshan & Janfeshan, 2021](#)) can improve language learning outcomes, there remains a gap in research regarding how AI and adaptive platforms influence learner beliefs. Learner beliefs about the effectiveness and

controllability of learning is crucial, as they directly relate to motivation, learning strategies, and attitudes (Horwitz, 1987; Ferreira Barcelos, 2015). AI systems capable of adapting materials, difficulty levels, and learning pace to individual learner profiles have the potential to foster positive beliefs, particularly the belief that language learning can be more personalized and responsive to individual needs.

Furthermore, the instant feedback provided by AI systems can reinforce the belief that mistakes are part of the learning process, in line with Vygotsky's (1978) zone of proximal development theory. This suggests that, in addition to enhancing cognitive achievement, technology can also shape affective aspects such as motivation and self-efficacy (Ustun et al., 2022). In the Indonesian context, however, EFL learner beliefs remain heavily influenced by exam orientation learning, leading students to prioritize vocabulary and grammar mastery over communicative competence (Suryanto, 2014). The integration of AI and adaptive platforms has the potential to shift these beliefs by providing instant feedback, personalized learning, and more communicative learning experiences, again resonating with Vygotsky's (1978) concept of the zone of proximal development. Nonetheless, challenges such as imitated digital infrastructure, inadequate internet access, and insufficient teacher readiness remain crucial determinants of how these technologies shape students' learning beliefs in Indonesia. Future research should explore more deeply how the integration of AI and adaptive learning platforms influences EFL learners' beliefs in Indonesia, particularly in shifting their orientation from mere exam preparation toward the development of communicative competence. Further studies should also investigate how AI and adaptive learning environments can transform learner beliefs to ensure that technology integration truly supports effective and inclusive language learning.

CONCLUSION

The learners' beliefs in acquiring EFL are contextual and shaped by the interaction between internal and external factors. Findings from various studies confirm that self-confidence, motivation, and socio-cultural norms play crucial roles in shaping language learning strategies and orientations. Thus, learners' beliefs are not static individual characteristics, but rather socially constructed understanding influenced by linguistic ideologies, cultural expectations, and specific learning experiences. This demonstrates the need for adaptive and context-sensitive pedagogical approaches that enable learners to develop positive beliefs supporting the growth of comprehensive language competence.

Furthermore, artificial intelligence (AI)-based technologies and adaptive learning platforms hold significant potential to foster more positive learner beliefs about the language learning process. Through instant feedback, personalized materials, and communicative learning experiences, these technologies can enhance motivation, self-efficacy, and the understanding that mistakes are an integral part of the learning. However, implementation within the Indonesian context still faces challenges such as limited

digital infrastructure, inadequate internet access, and insufficient teacher preparedness. Therefore, further research is needed to explore how AI integration can shift the learning orientations from test-focused preparation toward communicative competence development, ensuring that technology truly becomes a catalyst for forming learner beliefs that support effective and inclusive language learning.

ACKNOWLEDGEMENTS

The author extends sincere gratitude to all researchers whose work forms the basis of this review and to the editorial team for their professional guidance and support throughout the publication process.

REFERENCES

- Abdollahzadeh, E., & Rajaenia, M. (2024). English Language Learning Beliefs of Iranian High School Students. *TESL-EJ*, 28(1).
<https://doi.org/10.55593/ej.28109a6>
- Abdullah, K. N., & Razi, Ö. (2022). Investigation of Kurdish students' L2 motivational self-system and their motivational beliefs in high school. *Frontiers in Psychology*, 13.
<https://doi.org/10.3389/fpsyg.2022.974748>
- Al-Bogami, B., & Elyas, T. (2020). Promoting Middle School Students' Engagement Through Incorporating iPad Apps in EFL/ESL Classes. *SAGE Open*, 10(2).
<https://doi.org/10.1177/2158244020926570>
- Cai, Y., & Xing, K. (2023). Examining the mediation of engagement between self-efficacy and language achievement. *Journal of Multilingual and Multicultural Development*.
<https://doi.org/10.1080/01434632.2023.2217801>
- Casanova-Mata, I. (2023). Enhancing English Acquisition: Effects of among us Game-Based Gamification on Language Competence, Motivation, Attention, and Attitude towards the English Subject. *Education Sciences*, 13(11).
<https://doi.org/10.3390/educsci13111094>
- Dawala Wilang, J. (2021). Mindsets of High School Students in English Language Learning. *TESOL International Journal*, 16(2).
- Dornyei, Z., & Ryan, S. (2015). *The Psychology of The Language Learner Revisited (1st edition)*. Routledge.
<https://doi.org/10.4324/9781315779553>
- Ferreira Barcelos, A. M. (2015). Unveiling the relationship between language learning beliefs, emotions, and identities. *Studies in Second Language Learning and Teaching*, 5(2), 301–325.
<https://doi.org/10.14746/sslt.2015.5.2.6>
- Gabillon, Z. (2014). L2 Learner's Beliefs: An Overview. In *Journal of Language and Learning*, 3(2).
<https://www.researchgate.net/publication/259803455>
- Horwitz, E. (1987). *Surveying Students Beliefs about Language Learning*. In A. L. Wenden & J. Rubin (Eds.), *Learner Strategies in Language Learning* (pp. 119–129). NJ: Prentice-Hall.

- Huiyuan, G. (2024). Beliefs About and Motivation for Chinese EFL Learners' English Language Learning: A Study Drawing on the Theory of Planned Behaviour. *Asia Pacific Journal of Educators and Education*, 39(1), 25–44. <https://doi.org/10.21315/apjee2024.39.1.2>
- Janfeshan, K., & Janfeshan, M. M. (2021). The effect of otus social educational network on English achievement and attitudes of Iranian high school EFL learners. *Cogent Arts and Humanities*, 8(1). <https://doi.org/10.1080/23311983.2021.1895464>
- Kalaja, P., Barcelos, A. M. F., & Aro, M. (2017). Revisiting Research on L2 Learner Beliefs: Looking Back and Looking Forward. *The Routledge Handbook of Language Awareness*, 222–237. <https://doi.org/10.4324/9781315676494-14>
- Kalaja, P., Maria, A., Barcelos, F., & Aro, M. (2017). Revisiting Research on L2 Learner Beliefs Looking Back and Looking Forward. In P. Garrett & J. M. Cots (Eds.), *The Routledge Handbook of Language Awareness* (First Edition, pp. 222–237). Routledge.
- Kayyali, M. (2025). AI and Gamification: Engaging EFL Learners With Language Games. In Y. Baba Khouya & A. Ismaili Alaoui (Eds.), *Application of AI in the Teaching and Learning of English as a Foreign Language (EFL)*. IGI Global Scientific Publishing.
- Koné, K. (2024). An investigation into factors that shape Francophone African EFL learners' beliefs in learning English. *International Journal of Language Studies*, 18(3), 39–60. <https://doi.org/10.5281/zenodo.12604300>
- Mansouri, B., & Jami, P. Y. (2022). A Survey Study of Secondary Students' Beliefs and Preferences about Grammar Instruction: An Investigation of Iranian EFL Context. *MexTesol Journal*, 46(2), n2. <https://eric.ed.gov/?id=EJ1357789>
- Naghavian, M. (2024). English learners' beliefs about L2 speaking fluency: Insights from elicited metaphor analysis. *Modern Language Journal*, 108(3), 719–740. <https://doi.org/10.1111/modl.12952>
- Rabab'ah, G., Abuseileek, A., Abu Guba, M. N., & Bani Younes, M. A. (2025). Online Games and AI ChatGPT: 7th Grade English as a Foreign Language Learners' Vocabulary Acquisition and Preferences in the UAE. *Applied Research on English Language*, 14(2), 101–130. <https://doi.org/10.22108/arc.2025.143808.2420>
- Shirzad, S., Barjesteh, H., Dehqan, M., & Zare, M. (2022). Epistemic Beliefs and Learners' Self-Efficacy as Predictors of Language Learning Strategies: Toward Testing a Model. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.867560>
- Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333–339. <https://doi.org/10.1016/j.jbusres.2019.07.039>
- Suryanto, S. (2014). Issues in Teaching English in A Cultural Context: A Case of Indonesia. *Journal of English Literacy Education: The Teaching and Learning of English as a Foreign Language*, 2(1), 75–82. <https://doi.org/https://doi.org/10.36706/jele.v1i2.2075>
- Ustun, A. B., Simsek, E., Karaoglan-Yilmaz, F. G., & Yilmaz, R. (2022). The effects of AR-enhanced English Language Learning Experience on Students' Attitudes, Self-Efficacy and Motivation. *TechTrends*, 66(5), 798–809. <https://doi.org/10.1007/s11528-022-00757-2>
- Valeo, A., & Spada, N. (2016). Is there a better time to focus on form? Teacher and learner views. *Tesol Quarterly*, 50(2), 314–339. <https://doi.org/10.1002/tesq.222>
- Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. MA: Harvard University Press.
- Wallace, M. P., & Leong, E. I. L. (2020). Exploring language learning motivation among primary EFL learners. *Journal of Language Teaching and Research*, 11(2), 221–230. <https://doi.org/10.17507/jltr.1102.10>
- Waluyo, B. (2024). Technology-Fused English Teaching and Learning in Higher Education: From Individual Differences to Being Different Individuals. *Journal: Language Education and Acquisition Research Network*, 17(2), 42–47. <https://so04.tci-thaijo.org/index.php/LEARN/index>
- Zayed, N., & Razeq, A. H. A. (2021). Palestinian high school students' attitudes towards studying English language and culture. *International Journal of Arabic-English Studies*, 21(2), 7–28. <https://doi.org/10.33806/IJAES2000.21.2.1>

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright ©2025 Vega Hesmatantya. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Enhancing EFL students' reading comprehension through Instagram: A quasi-experimental study

Sudiran*,¹ Wahyu Taufiq,² A.B. Prabowo Kusumo Adi³

¹Universitas Muhammadiyah Malang, Indonesia, ²Universitas Muhammadiyah Sidoarjo, Indonesia,

³Universitas PGRI Semarang, Indonesia

The rapid development of digital technology has reshaped students' learning preferences, particularly in accessing reading materials. Despite the abundance of digital resources, many students still struggle to develop sufficient English reading skills, indicating a need for innovative media that align with their learning habits. While several studies have investigated social media as a supplementary learning tool, limited research has specifically examined the systematic use of Instagram in enhancing students' reading comprehension. Therefore, this study aimed to examine the implementation of Instagram as a learning medium to improve students' reading skills and to measure its effectiveness in classroom practice. A quasi-experimental design was employed with 28 university students, who were taught with Instagram-based reading materials. Reading materials were selected from authentic Instagram posts and delivered during eight instructional sessions over eight weeks. Data were collected through validated and reliable reading comprehension tests administered as a pretest and post-test. In addition, a student reading test was used to capture learners' skills in Instagram-based reading activities. The findings revealed that students' post-test scores significantly outperformed their pre-test scores, with a notable increase in their mean post-test scores ($M = 69.00$) compared to the pretest scores ($M = 59.60$). Independent samples T-test analysis confirmed a statistically significant difference ($p < 0.05$), indicating that Instagram effectively supports reading comprehension improvement. Moreover, students reported higher motivation and engagement when learning through Instagram. These results suggest that Instagram can serve as an innovative and authentic medium to enhance reading skills in EFL contexts.

Keywords: Implementation, Instagram, Improving, Reading, Students

OPEN ACCESS

ISSN 2503 3492 (online)

*Correspondence:

Sudiran

sudiran@gmail.com

Received: 12th July 2025

Accepted: 23th September 2025

Published: 31th October 2025

Citation:

Sudiran, S., Taufiq, W., Adi, A.B.P.K. (2025). Enhancing EFL students' reading comprehension through Instagram: A quasi-experimental study. *JEES (Journal of English Educators Society)*, 10(2). <https://doi.org/10.21070/jees.v10i2.1957>

INTRODUCTION

The use of social media platforms such as Instagram has the potential to improve language learning, with a more interactive and student-centered approach. Instagram can significantly improve writing skills and intrinsic motivation (Putri et al., 2025; Rahayu et al., 2025). The use of social media and support from educational institutions, can enrich students' learning experiences and encourage their active involvement (Hadinuddin et al., 2025).

Mobile Assisted Language Learning is prevalent in teaching English, but its effectiveness and practicality are rarely well confirmed (Mulyadi et al., 2022). Mobile learning media can be used as student-centered language learning (Cahyana et al., 2018; Burston, 2017). Research findings encourage teachers to increase the popularity of Instagram among students for learning purposes through classroom learning (Damaryanan & Subekti, 2024).

Improving reading skills requires learning media that are familiar to students. Mobile learning media, such as smartphones and Instagram, are considered more accessible and practical both inside and outside the classroom compared to computer-based learning media because most students already have smartphones ([Mutiarasari et al., 2020](#); [Khasanah & Rizal, 2023](#)). Learning English using mobile phones or Instagram allows students to access learning resources more quickly ([Wan Azli et al., 2018](#)).

Most students today feel that reading is not an interesting activity because they get bored easily and pay less attention to reading skills. The commitment they gave to reading activities before implementing Instagram videos was quite low ([Morshidi et al., 2019](#); [Khairuddin, 2013](#)). Students have a good perception about using Instagram for language learning because the videos and photos uploaded on Instagram accounts are easy to understand and fun ([Astiyandha & Irwansyah, 2024](#)). Students are not aware of how their interests affect their comprehension abilities, so it is important to find this relationship ([O'Flynn, 2016](#)). Instagram can be utilized as a teaching tool in reading classes because of its effectiveness and functionality ([Rinda et al., 2018](#); [Niño et al., 2024](#); [Syachsalsabillah & Hamid, 2024](#)).

Reading is the ability to understand a text and to summarize it well ([Reflianto et al., 2021](#); [Azmuddin et al., 2020](#)). Reading tests are used to improve students' reading skills. The discourse test must contain information that demands student understanding ([Irwansyah et al., 2019](#); [Febrina et al., 2019](#)). The reading ability test can connect and generalize discourse concepts through synthesis skills. These skills consist of synthetic-level cognitive activities as a high-level and complex activity ([Lazarus, 2020](#)).

Students consider Instagram as a valuable tool for interactive, interesting, and enjoyable English learning as a major factor in developing their positive learning experience ([Saifuddin et al., 2024](#); [Sarwoko et al., 2024](#); [Thomas & Park, 2020](#)). Inspired by literature that has shown the benefits of using Instagram as a teaching and learning tool ([Handayani, 2015](#); [Khalitova & Gimaletdinova, 2016](#)), researchers designed Instagram reading activities to measure student reading ability improvement. Through Instagram, teachers can create a dynamic and interactive learning environment for students ([Méndez, 2024](#); [Handayani et al., 2018](#)).

The use of Instagram encourages students to have a positive attitude towards the use of mobile technology in learning new vocabulary ([Alzahrani, 2015](#); [Zárate & Cisterna, 2017](#); [La'biran et al., 2024](#); [Baruti & Subekti, 2023](#); [Al-Ali, 2014](#)). Instagram allows students to exchange views and opinions on various topics ([Suryantari & Priyana, 2018](#); [Akhiar et al., 2017](#)) They read each other's writings through the use of Instagram ([Thomas, 2019](#); [Ramalia, 2021](#); [Alhabash & Ma, 2017](#)).

The research on the Enhancing EFL Students' Reading Comprehension through Instagram is important to conduct because it is still limited, and there are still few researchers who have conducted research on this theme. Meanwhile, this study aims to reveal the students' reading comprehension improvement and to explore students' reading ability through Instagram.

The difference between the results of this study and previous studies is that this study used a quasi-experiment by comparing the pretest and posttest scores of students' reading abilities. The results of data analysis show that Instagram plays an important role in improving learning outcomes in reading courses. Instagram, as an interpretive reading learning medium, can increase student learning motivation because this platform is interesting and fun. It is an application for accessing English vocabulary, for searching for reading texts, and for sharing information.

Despite a growing body of research on the use of Instagram in English language teaching, most previous studies have focused more on student motivation and engagement in improving writing skills than on reading comprehension skills. Furthermore, previous studies often prioritized student perceptions of Instagram use, employing survey-based designs, with limited experimental evidence examining the effect of Instagram use on student reading performance. Another limitation is that many studies used teacher-designed, textbook-based materials, rather than authentic social media content, which could reflect real-world English language skills and improve students' reading ability.

To address this gap, this study investigates the implementation of Instagram as a medium to improve EFL students' reading skills using authentic Instagram texts, visuals, and comments in a quasi-experimental design. The novelty of this study lies in its dual focus on the cognitive and affective dimensions of reading. Furthermore, by situating this research in the Indonesian EFL context, where reading skills are often perceived as difficult and tedious, this study provides theoretical insights and practical implications for integrating Instagram into higher education language learning.

METHODS

This study employed a quasi-experimental design with a one-group pretest–post-test approach to examine the effectiveness of Instagram-based instruction in improving students' reading comprehension. A pretest was administered before the intervention, followed by an eight-week treatment, and concluded with a post-test. To enhance methodological rigor, validity, and reliability checks of the instruments, treatment fidelity, and statistical analyses were carefully conducted.

Participants

The population of this study comprised 150 undergraduate students (six intact classes) enrolled in the fourth semester of the English Education Study Program at the University of Muhammadiyah Malang during the 2024/2025 academic year. Using purposive sampling, one intact class of students enrolled in the *Reading Comprehension* course (n = 28) was selected as the sample.

The sample consisted of 28 students (22 female, 6 male), aged 19–21 years. Participants reported daily use of Instagram, averaging 1–2 hours, ensuring familiarity with the platform. The research sample was determined based on the following criteria: (a) students were in their 4th semester, (b) they were officially registered in the course, and (c) they were active Instagram users with internet access.

Data collection technique and procedure

This study used a Reading Comprehension Test as data. A 40-item multiple-choice test was developed to assess students' reading comprehension. The test items were classified into three skill domains: literal comprehension (16 items), inferential comprehension (14 items), and contextual vocabulary (10 items). The text sources included authentic Instagram captions, short narratives, expository texts, and argumentative paragraphs relevant to the treatment. Each item had four choices with one correct answer (scoring: 1 = true, 0 = false; maximum score = 40).

To ensure the quality of the research instrument, several procedures were undertaken to establish its validity and reliability. Content validity was verified by three experts in English education and reading pedagogy, who carefully reviewed each test item for its relevance, clarity, and representativeness to the intended reading skills. The experts' evaluations yielded a Scale-level Content Validity Index (S-CVI/Ave) of 0.92, which exceeds the recommended minimum threshold of 0.90, indicating that the items were highly valid in measuring the targeted constructs. Furthermore, pilot testing was conducted with 30 non-participant students to examine the difficulty and discrimination indices of the test items. Only items that met acceptable statistical criteria difficulty indices ranging from 0.30 to 0.80 and discrimination indices above 0.30 were retained for the final test. Finally, the reliability of the instrument was determined using the Kuder–Richardson Formula 20 (KR-20), which produced a coefficient value of 0.87, demonstrating high internal consistency and confirming that the test items were stable and reliable for assessing students' reading comprehension.

The study was carried out over a ten-week period, following a structured sequence to ensure systematic implementation and consistency across all research stages. In the **first week**, students completed a 40-item reading comprehension pretest administered under standardized paper-based conditions within a 60-minute time limit. This pretest was designed to assess their initial reading ability prior to the intervention. The **treatment phase**, conducted over **eight consecutive weeks (Weeks 2–9)**, involved weekly instructional sessions lasting 90 minutes each. During these sessions, students engaged in a variety of Instagram-based reading activities, including analyzing captions, interpreting infographic posts, responding to persuasive messages, and summarizing short videos. All instructional sessions adhered to a detailed lesson plan to maintain treatment fidelity and ensure that learning objectives were consistently met. Finally, in the **tenth week**, students completed a post-test that was equivalent and parallel in format, difficulty, and content coverage to the pretest, thereby minimizing potential testing effects and ensuring comparability of results. This sequence of activities enabled the researchers to systematically measure the impact of Instagram-based learning on students' reading comprehension.

To monitor treatment fidelity, the instructor used a checklist for each session, ensuring that all planned activities were implemented consistently. Engagement was tracked by recording students' number of posts, comments, and task submissions on Instagram. Test papers were scored

dichotomously (correct = 1, incorrect = 0). Raw scores were converted into percentages and categorized as follows: Very Good ($\geq 85\%$), Good (70–84%), Moderate (55–69%), Low (40–54%), and Very Low ($< 40\%$). Data were entered into SPSS v.25 for statistical analysis.

Data Analysis

Data analysis in this study followed a systematic quantitative procedure to examine the effectiveness of Instagram-based instruction on students' reading comprehension. Initially, **descriptive statistics** including the mean, standard deviation, and frequency distribution were calculated to summarize students' pretest and post-test performance. The **normality of gain scores** was then tested using the **Shapiro–Wilk test** to determine the appropriate statistical method for further analysis. For data that met the assumption of normality, a **paired-samples t-test** was conducted to identify significant differences between pretest and post-test mean scores.


Conversely, when the data did not meet normality assumptions, a **Wilcoxon signed-rank test** served as the non-parametric alternative. To assess the magnitude of improvement, **effect sizes** were computed using **Cohen's d** for parametric data and r for non-parametric data. In addition, **Pearson's correlation analysis** was performed to explore the relationship between students' engagement metrics on Instagram such as the number of posts and comments and their reading comprehension gains. This comprehensive analytical approach allowed for both statistical verification of improvement and exploration of engagement-related factors contributing to students' reading development.

RESULTS AND DISCUSSION

The research was conducted for eight weeks in the Interpretive Reading class of the English Language Education study program, FKIP, University of Muhammadiyah Malang. The first week is used to give a pre-test to 28 students. From the pre-test activities, an average score of 59.60 was obtained. After the pre-test activities were carried out, students were then given lecture material adapted from Instagram content. There are several Instagram contents that are used as learning media as well as lecture material. The content is accessed via the internet and then selected and delivered to students as learning media in class. The Instagram content is delivered to students for 6 X 100 minutes according to the lecture schedule. Class schedules with Instagram content as learning media can be seen in [Table 1](#) as follows.

After presenting lecture material for eight weeks, accompanied by media from Instagram content as a treatment, the next activity is measuring student learning outcomes. Measurement of student learning outcomes is done by giving a post-test or final test. This final test was designed not only to measure learning outcomes but also to see the role of Instagram media in students' reading skills. The final test was prepared using an objective test model with multiple-choice forms. The use of objective tests is adjusted to the competency being measured, namely, interpretive reading. As a reading skill, interpretive reading requires an instrument that can measure student achievement validly.

TABLE 1 | Instagram Content as Learning Media

No.	Day	Instagram Content	Message Content
1	19-09-2022	Pre-test	Score: 59.60
2	26-09-2022		<p>Sayap-Sayap Patah Movie Content. Sayap-Sayap Patah is a film that tells the story of the romance between a man and a woman, played by Aji, who works as a police officer, and his wife, Nani. The story of the Sayap-Sayap Patah Film was adapted from the true story of the 2018 Mako Brimob Riot.</p>
3	03-10-2022		<p>Citayem Fashion Show content. Citayem Fashion Show is an arena for creativity in appearance by teenagers or newly grown children. Citayem Fashion Show activities are widely discussed by people, generally netizens on social media. Their activities include taking photos or selfies around Jakarta, such as Jalan Jenderal Sudirman, Citayem, Bojonggede, and Depok (SCBD). They show off the clothes and accessories they wear by waddling in the traffic light road markings when the vehicle stops waiting for a red light. Their style is like the models walking on the catwalk, so that many road users look up to them. Not infrequently, the road users stop and pay attention to them because the appearance of the teenager is attractive and attracts attention.</p>
4	10-10-2022		<p>Culinary Content. Various kinds of cuisine in the city of Malang. Gebuk Sate Shop provides a menu of beef satay, which is crushed so that the meat is soft. This dish is very popular among Malang residents because of its delicious taste.</p>

5 17-10-2022



Food and beverage content. Special food from Malang, grilled meatballs, and squid chicken noodles. This shop serves a food menu that is different from those in other cities. This shop provides food such as chicken, squid noodles, and grilled meatballs, as well as various drinks such as young coconut ice, ginger tea, mixed ice, etc.

6 24-10-2022



Entertainment Content. The Beachbuds is an animation loved by children in countries such as Australia, Japan, and Latin America. This entertainment is broadcast on television and gets a wide response from its viewers. The Beachbuds is broadcast in prime time so that it gets a high rating like other animations, namely, SpongeBob, SquarePants, and Lego.

7 31-10-2022



Environmental Content. This content tells us all about the dangers of tobacco, which can kill humans and poison the earth. This video clip reminds us to leave tobacco in cigarettes because it damages the body through toxic substances in them, such as Nicotine, Tar, Benzene, etc.

8 07-11-2022 post-test

Score: 69.00

It is evident from the final exam results that Instagram media enhances students' comprehension of the course subject. The final test results of 28 students obtained an average score of 69.00. When compared with the pre-test value of 59.60, the post-test value is greater. The difference between the pre-test and post-test is 9.4. If the difference in

pre-test and post-test scores is analyzed using the SPSS program, the result is a T-test = 0.962. The results of this T-test are significant because the p-value is greater than the Sig value. 0.000. The relationship between the Pre-test, Treatment, and Post-test scores can be seen in [Table 2](#) below as follows.

TABLE 2 | Pre-test, Treatment, and Post-test

Pre-test Score	Treatment	Post-test Score
59.60	8 Meeting/weeks	69.00

According to statistical data, there is a correlation between before and after the learning process using Instagram as a learning medium. This data can be seen in [Table 3](#) below.

TABLE 3 | Statistical Data Analysis Paired Samples Statistics

	Mean	N	Std. Deviation	Std. Error Mean
Pair 1 Sebelum Belajar	59.6071	28	2.85889	.54028
Setelah Belajar	69.0000	28	2.82843	.53452

[Table 3](#) above explains that before giving the material using Instagram media, the average test value was 59.60, the number of students was 28, the standard deviation was 2.85889, and the mean standard error was 0.54028. Meanwhile, the score after giving the material using Instagram media was the average value; the test mean was 69.00, the number of students was 28, the standard deviation was 2.82843, and the mean standard error was 0.53452. This means that there is a positive correlation between before and after the learning process using Instagram media.

Meanwhile, the significance value of the T-test for pre-test and post-test values can be seen in [Table 4](#) below.

TABLE 4 | The Significance of the T-test Paired Samples Correlations

	N	Correlation	Sig.
Pair 1 Sebelum Belajar & Setelah Belajar	28	.962	.000

[Table 4](#) explains that the Sig. $0.000 < 0.05$. The correlation value is 0.926. It can be concluded that there is a significant relationship between before and after learning with Instagram media.

The study indicates that Instagram can be used as a learning medium to improve student reading achievement. It has been proven that there are significant differences between before and after learning using Instagram media. Before learning, the pre-test results were 59.60. After learning 6 times face-to-face/week using Instagram learning media, the post-test results show the number 69.00. The results of statistical analysis using SPSS show that the T-test results are 0.962. This means that there is a significant difference between before and after teaching and learning activities using Instagram media. Instagram is a smartphone application for learning English, such as searching for reading texts and vocabulary. The results of the research answered the research problems previously mentioned, namely: What is the role of Instagram in improving students' reading skills?

The results of this study were supported by data consisting of 28 students as the research sample. From the research results, it can be said that using Instagram helps students develop reading skills. Using Instagram can help them find interesting ideas from reading texts. Instagram can be used to help them improve their English vocabulary. This means that applying Instagram as a learning medium in English classes can help students improve their reading skills, access valuable ideas, and improve vocabulary mastery.

The results of this quasi-experimental study indicate that Instagram-based learning significantly improves EFL students' reading comprehension. The improvement in post-test scores compared to pre-test results suggests that integrating authentic Instagram materials into reading instruction can increase engagement, motivation, and comprehension. These results are consistent with previous research that observed the positive influence of social media-based learning on students' linguistic and affective development ([Putri et al., 2025](#); [Rahayu et al., 2025](#)). However, this study extends previous research by providing experimental evidence, rather than solely relying on self-reported perceptions.

The observed improvements can be attributed to several pedagogical efforts on Instagram. First, Instagram's multimodal nature, which combines text, image, and comment interactions, supports students' cognitive processing and contextual understanding. Based on multimedia learning theory, dual-channel input that is visual and verbal enhances comprehension by facilitating more concrete encoding and integration of meaning. Furthermore, [Al-Ali \(2014\)](#) and [Akhiar et al. \(2017\)](#) explain that image-supported text helps students read critically and imaginatively. In this study, students engaged in reading activities involving Instagram captions, short narratives, and infographic posts, all of which improved inferential and interpretive comprehension skills. These activities align with constructivist learning theory, which explains that students construct meaning through active interaction with authentic and meaningful content.

Second, enhanced reading performance can also be related to affective engagement. As previously reported, enjoyment, curiosity, and relevance significantly support reading motivation in digital contexts ([Méndez, 2024](#); [Thomas, 2019](#)). The familiar and interactive environment of Instagram lowers learner anxiety and increases positive attitudes toward reading ([Sarwoko et al., 2024](#)). These findings support the notion that students' positive emotional connection with digital media increases their willingness to read and comprehend English texts ([Khalitova & Gimaletdinova, 2016](#)). Therefore, affective engagement should be considered a key mediating factor between Instagram use and reading comprehension achievement.

Third, the findings indicated that Instagram serves as a form of Mobile Assisted Language Learning (MALL) that encourages self-directed learning. Students' daily exposure to English through Instagram creates incidental learning opportunities that foster vocabulary retention and contextual understanding ([Burstson, 2017](#); [Niño et al., 2024](#)). The integration of Instagram into classroom activities facilitates both formal and informal learning processes, aligning with the

blended learning paradigm. Therefore, the quasi-experimental results support the argument that mobile applications effectively complement traditional reading instruction, especially in EFL contexts where exposure to authentic English materials is limited (Mutiarasari et al., 2020).

Despite the positive results, the score increase (from 59.60 to 69.00) was moderate. This moderate increase suggests that while Instagram increased engagement and comprehension, its effects may have been limited by the short duration of the intervention and the limited scope of the reading tasks. Short-term interventions develop early-level comprehension (literal comprehension), while higher-order skills such as inference making, synthesis, and critical evaluation require longer exposure and structured instruction. Furthermore, the use of multiple-choice tests, while reliable, limits the interpretive and analytical dimensions of reading comprehension. Future research could employ a mixed-methods approach that combines quantitative testing and qualitative discourse analysis to capture a wider range of learning outcomes.

The present study contributes new evidence to the growing body of research on social media in English as a Foreign Language (EFL) education by changing the focus from writing and motivation to reading comprehension outcomes. While previous studies (e.g., Damaryanan & Subekti, 2024; Syachsalsabillah & Hamid, 2024) have primarily focused on Instagram's motivational effects, this study empirically demonstrates its cognitive benefits through a quasi-experimental framework. The findings thus support a more holistic understanding of digital literacy, in which the cognitive (comprehension) and affective (interest and motivation) domains are interconnected.

From an educational perspective, this study confirms that English language instructors need to design reading activities that utilize authentic and multimodal Instagram content. Tasks such as text translation, comment analysis, and thematic summaries help students connect visual stimuli with textual meaning, thereby enhancing comprehension. Furthermore, instructors need to guide students in critically evaluating online content, which enhances digital literacy and critical reading, thus developing skills essential for 21st-century communication. Institutional support plays a crucial role in providing digital infrastructure and teacher training to effectively integrate mobile technology in EFL classrooms.

Finally, although this study is limited by its small sample size and focus on a single institution, it provides a valuable foundation for future comparative research. Longitudinal or cross-platform studies (e.g., comparing Instagram with TikTok, YouTube, or Vlogs) could further explain how different digital uses affect specific aspects of reading ability. Overall, this study provides empirical support for the use of Instagram in higher education EFL contexts, indicating its potential to enhance reading comprehension through authentic, multimodal, and engaging learning experiences.

CONCLUSION

Based on the research results, it can be concluded that Instagram can be used as a learning medium to improve

reading skills. This means integrating Instagram in the classroom allows students to improve learning outcomes. This is because Instagram can help them access the main ideas and vocabulary of a reading text.

This study demonstrated that the integration of Instagram-based materials into EFL reading instruction can positively enhance undergraduate students' reading comprehension skills. The quasi-experimental design revealed a statistically significant improvement in post-test performance compared to the pre-test, indicating that Instagram content provides meaningful support in developing interpretive reading. Although the observed improvement was modest, the findings confirm that digital platforms can serve as effective supplementary tools when embedded within structured pedagogical frameworks.

From a pedagogical perspective, these results suggest that English lecturers should consider incorporating Instagram into reading courses as an authentic and multimodal resource. By engaging students with real-world content that ranges from cultural, culinary, and entertainment texts to environmental issues, teachers can foster not only comprehension skills but also vocabulary development and critical engagement with diverse genres. Importantly, Instagram can be used to design collaborative classroom activities, stimulate discussions, and motivate learners by bridging academic reading practices with students' digital literacy practices. However, instructors are encouraged to integrate Instagram alongside conventional materials and guided instructional strategies, ensuring that its use supports rather than replaces systematic reading instruction.

For future research, several avenues deserve attention. First, longitudinal studies are needed to examine the sustained impact of Instagram on higher-order reading skills, including critical analysis and inferencing. Second, comparative research with other digital platforms such as TikTok, YouTube, or digital newspapers could illuminate the platform-specific affordances for literacy development. Third, future studies should expand to larger and more diverse samples to increase generalizability. Finally, adopting mixed-methods designs, which combine test-based data with qualitative measures such as discourse analysis, learner reflections, or classroom interaction analysis, would provide richer insights into how Instagram shapes students' reading processes and engagement.

In sum, this study contributes empirical evidence to the growing scholarship on digital literacy in EFL contexts, underscoring the pedagogical potential of Instagram as an accessible and motivating tool for enhancing reading comprehension. By situating language learning within students' everyday digital environments, educators can create more contextualized, engaging, and effective reading instruction in higher education.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the generous financial support of the University of Muhammadiyah Malang, Indonesia. Furthermore, the authors declare that artificial intelligence tools such as Google Translate, ChatGPT, and

Grammarly were used for language editing.

REFERENCES

- Akhiar, A., Mydin, A. A., & Kasuma, S. A. A. (2017). Students' perceptions and attitudes towards the use of Instagram in English language writing. *Malaysian Journal of Learning and Instruction, Special Issue*, 47–72. <https://doi.org/10.32890/mjli.2017.7796>
- Al-Ali, S. (2014). Embracing the Selfie Craze: Exploring the Possible Use of Instagram as a Language mLearning Tool. *Issues and Trends in Learning Technologies*, 2(2), 1–17. https://doi.org/10.2458/azu_itet_v2i2_ai-ali
- Alhabash, S., & Ma, M. (2017). A tale of four platforms: Motivations and uses of Facebook, Twitter, Instagram, and Snapchat among college students?. *Social media+ society*, 3(1), 2056305117691544. <https://doi.org/10.1177/2056305117691544>
- Alzahrani, H. (2015). Examining the Effectiveness of Utilizing Mobile Technology in Vocabulary Development for Language Learners. *Arab World English Journal (AWEJ)*, 6(3), 108–119. <https://dx.doi.org/10.24093/awej/vol6no3.7>
- Astiyandha, T., & Irwansyah, M. L. (2024). Analysis the college students' perception on aarons.english instagram account in speaking. *Lingua*, 20(1), 83–104. <https://doi.org/10.34005/lingua.v20i1.3713>
- Azmuddin, R. A., Nor, N. F. M., & Hamat, A. (2020). Facilitating online reading comprehension in enhanced learning environment using digital annotation tools. *IAFOR Journal of Education*, 8(2), 7–27. <https://doi.org/10.22492/ije.8.2.01>
- Baruti, T. D. W. P., & Subekti, A. S. (2023). Instagram to learn English vocabulary: a study of Indonesian non-English major university students. *IJEE (Indonesian Journal of English Education)*, 10(1), 106–125. <https://doi.org/10.15408/ijee.v10i1.26969>
- Burston, J. (2017). MALL: Global prospects and local implementation. *Call-Ej*, 18(1), 1–8. <https://callej.org/index.php/journal/article/view/231>
- Cahyana, U., Paristiowati, M., & Fauziyah, S. (2018). Development of Android-Based Mobile learning media on Atomic Structure and Periodic Table. *IOP Conference Series: Materials Science and Engineering*, 434, 1–7. <https://doi.org/10.1088/1757-899X/434/1/012095>
- Damaryanan, R. E., & Subekti, A. S. (2024). Indonesian Non-English Major Learners' Attitudes Toward Instagram to Practice Speaking in English: A Survey Study. *Acuity: Journal of English Language Pedagogy, Literature, and Culture*, 9(1). <https://doi.org/10.35974/acuity.v9i1.3240>
- Febrina, Bustami, U., & Asnawi, M. (2019). Analysis of Reading Comprehension Questions by Using Revised Bloom's Taxonomy on Higher Order. *English Education Journal*, 10(1), 1–15.
- Hadinuiddin, M., Mansur, H., & Salim, A. (2025). Systematic Literature Review: The Utilization of Social Media in Learning to Increase Student Engagement. *Indonesian Journal of Social Technology*, 6(1), 155–168. <http://jist.publikasiindonesia.id/>
- Handayani, A. D., Cahyono, B. Y., & Widiati, U. (2018). The Use of Instagram in the Teaching of EFL Writing: Effect on Writing Ability and Students' Perceptions. *Studies in English Language Teaching*, 6(2), 112. <https://doi.org/10.22158/selt.v6n2p112>
- Handayani, F. (2015). Instagram as a teaching tool? Really? *Proceedings of ISELT FBS Universitas Negeri Padang*, 4(1), 320–327.
- Irwansyah, D., Nurgiyantoro, B., & Sugirin. (2019). A literature-based reading instructional model for an Islam-affiliated University in Indonesia. *International Journal of Instruction*, 12(3), 577–594. <https://doi.org/10.29333/iji.2019.12335a>
- Khairuddin, Z. (2013). A Study of Students' Reading Interests in a Second Language. *International Education Studies*, 6(11), 160–170. <https://doi.org/10.5539/ies.v6n11p160>
- Khalitova, L., & Gimaletdinova, G. (2016). Mobile Technologies in Teaching English as a Foreign Language in Higher Education: a Case Study of Using Mobile Application Instagram. *ICERI2016 Proceedings*, 1(March 2017), 6155–6161. <https://doi.org/10.21125/iceri.2016.0395>
- Khasanah, N., & Rizal, D. (2023). The use of English learning contents on instagram for improving English speaking skills among EFL students' perception. *ELT Echo: The Journal of English Language Teaching in Foreign Language Context*, 8(2), 211. <https://doi.org/10.24235/eltecho.v8i2.15620>
- La'biran, R., Triana Ruung, T., & Sallata, Y. N. (2024). Exploring the Influence of Instagram Usage on English Education Students' Vocabulary Acquisition at UKI Toraja. *International Journal of Social Science and Human Research*, 07(01). <https://doi.org/10.47191/ijsshr/v7-i01-62>
- Lazarus, K. U. (2020). Socio-demographic factors affecting reading comprehension achievement among secondary school students with learning disabilities in Ibadan, Nigeria. *IAFOR Journal of Education*, 8(1), 145–157. <https://doi.org/10.22492/ije.8.1.09>
- Méndez, N. C. (2024). Instagram's Impact on Rural Colombian English Education: A Fourth Industrial Revolution Perspective. *Journal of Digital Learning and Education*, 4(2), 100–113. <https://doi.org/10.52562/jdle.v4i2.1056>
- Morshidi, A., Embi, M. A., & Hahim, H. (2019). Instagram Application: An Active Tool in Cultivating Reading Behaviour. *Journal of Information Systems and Technology Management*, 4(11), 95–106. <https://gaexcellence.com/jistm/article/view/1065/913>
- Mulyadi, D., Aimah, S., Arifani, Y., & Singh, C. K. S. (2022). Boosting EFL Learners' Listening Comprehension through a Developed Mobile Learning Application: Effectiveness and Practicality. *Applied Research on English Language*, 11(3), 37–56. <https://doi.org/10.22108/are.2022.130726.1785>

- Mutiarasari, M. I., Ngadiso, & Kristina, D. (2020). Indonesian learners' perception on the usage of mobile applications in writing skills. *Leksika*, 14(1), 36–42. <https://doi.org/10.30595/lks.v14i1.5697>
- Niño, C., Portilla, R., Vega, S., Natalia Cardozo Niño, H., Paola Rojas Portilla, E., & Marcela Sierra Vega, C. (2024). Implementation Of Instagram to Foster English Lexical Competence Development: An Action Research. *Bilingual Research Academic Journal*, 1(1), 16–35. <https://doi.org/10.24054/braj.v1i1.3177>
- O'Flynn, K. (2016). Students' Reading Interests Impact on Reading Comprehension Abilities How has open access to Fisher Digital Publications benefited you? Students' Reading Interests Impact on Reading Comprehension Abilities. St. John Fisher College.
- Putri, N., Saehu, A., & Bactiar. (2025). Boosting English Motivation and Writing Proficiency through Instagram Captions in Discovery-Based Learning. *Indonesian Journal of Educational Science*, 7(2). <https://doi.org/10.31605/ijes.v7i2.4842>
- Rahayu, N. S., Rahayu, Y. S., & Ramadhani, R. F. (2025). Unveiling the Effectiveness of Instagram to enhance Descriptive Text writing. *English Education and Literature Journal*, 5(2), 171–180. <https://doi.org/10.53863/e-jou.v5i02.1527>
- Ramalia, T. (2021). The Students' Perspective of Using Instagram as a Writing Assignment Platform. *J-SHMIC: Journal of English for Academic*, 8(2), 122–131. [https://doi.org/10.25299/jshmic.2021.vol8\(2\).7433](https://doi.org/10.25299/jshmic.2021.vol8(2).7433)
- Reflianto, Setyosari, P., Kuswandi, D., & Widiati, U. (2021). Reading comprehension skills: The effect of online flipped classroom learning and student engagement during the COVID-19 pandemic. *European Journal of Educational Research*, 10(4), 1613–1624. <https://doi.org/10.12973/EU-JER.10.4.1613>
- Rinda, R. K., Novawan, A., & Miqawati, A. H. (2018). Students' perspectives on social media-based learning of writing through Instagram. *Journal of English in Academic and Professional Communication*, 5(1). <https://publikasi.polije.ac.id/jeapco/article/view/1261/854>
- Saifuddin, M., Islam, M. H., & Hasanah, S. (2024). Students' Perception of Utilizing Instagram as a Learning Aid in Their Speaking English. *Journal of English Language Education and Literature*, 9(1), 40–50.
- Sarwoko, G. P., Sabudu, D., & Rorimpandey, R. (2024). Students' Perception Towards Instagram as English Mobile learning. *Journal of Teaching English, Linguistics, and Literature*, 3(11), 1450–1476.
- Suryantari, H., & Priyana, J. (2018). Exploring Ways of Using Facebook and Instagram in Teaching English. In *International Conference of Communication Science Research (ICCSR 2018)*, 298–301. Atlantis Press. <https://doi.org/10.2991/iccsr-18.2018.66>
- Syachsalsabillah, S., & Hamid, A. I. R. (2024). English Education Students' Perception and Motivation in Learning Descriptive Writing Skills Through Instagram. *Journal of English Language Learning (JELL)*, 8(1), 506–517. <https://doi.org/10.31949/jell.v8i1.9758>
- Thomas, K. (2019). Reading and Writing Activities on Instagram. *JALT Postconference Publication*, 472–481. <https://doi.org/10.37546/jaltpcp2019-54>
- Thomas, K., & Park, S. (2020). Integrating Instagram into the Curriculum of a Japanese Freshmen English Communication Course for Reading and Writing Activities Kevin Thomas and Simon Park, Asia University. *CELE Journal*, 28, 93–116. <https://doi.org/11300289>
- Wan Azli, W. U. A., Shah, P. M., & Mohamad, M. (2018). Perception on the Usage of Mobile Assisted Language Learning (MALL) in English as a Second Language (ESL) Learning among Vocational College Students. *Creative Education*, 09, 84–98. <https://doi.org/10.4236/ce.2018.91008>
- Zárate, P., & Cisterna, C. (2017). Action Research: The Use of Instagram as an Interactive Tool for Developing the Writing of Short Stories. *European Journal of Education Studies*, 2(8), 527–543. <https://doi.org/10.5281/zenodo.1035497>

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright ©2025 Sudiran, Wahyu Taufiq, A.B. Prabowo Kusumo Adi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Enhancing students' speaking skills using ChatGPT application in ESP maritime English learning

Dyah Ratnaningsih*

Politeknik Pelayaran Surabaya, Indonesia

This research investigated the impact of utilizing ChatGPT, an advanced artificial intelligence language model, on improving maritime English-speaking skills among students at a maritime polytechnic in Indonesia. Effective communication in maritime operations is critical for maintaining safety, preventing misunderstandings, and ensuring efficient coordination among multinational crew members. However, many maritime students face challenges in attaining adequate English-speaking proficiency necessary for these operations. This study adopted a classroom action research (CAR) design involving 69 students over two instructional cycles. Data collection methods included pretests and post-tests to quantitatively measured speaking performance, alongside questionnaires to gather qualitative insights into students' perceptions of the learning process and their engagement. Quantitative data were analyzed using paired sample t-tests in SPSS 22, revealing statistically significant improvements in key areas such as vocabulary use, grammatical accuracy, and overall fluency after the introduction of ChatGPT-assisted learning activities. Additionally, qualitative findings from questionnaires indicated enhanced student confidence, motivation, and active participation in speaking tasks, suggesting positive attitudes toward this AI-supported learning approach. The study highlighted the potential of integrating ChatGPT as a supplementary instructional tool within maritime English curricula to support language acquisition more effectively. By fostering maritime English-speaking competence through innovative technology integration, this approach contributed to better preparation of students for real-world communication demands in the maritime industry. Ultimately, the findings demonstrated that ChatGPT can serve as a valuable resource for language instructors, promoting more engaging, interactive, and productive learning environments that align with the operational needs of global maritime professionals.

Keywords: ChatGPT, ESP, maritime English, speaking

OPEN ACCESS

ISSN 2503 3492 (online)

*Correspondence:

Dyah Ratnaningsih

dvalifhana@poltekpel-sby.ac.id

Received: 1st August 2025

Accepted: 30th August 2025

Published: 31th October 2025

Citation:

Ratnaningsih, D. (2025). Enhancing students' speaking skills using ChatGPT application in ESP maritime English learning. *JEES (Journal of English Educators Society)*, 10(2). <https://doi.org/10.21070/jees.v10i2.1964>

INTRODUCTION

The maritime industry relies fundamentally on effective communication to ensure safety, operational efficiency, and adherence to international regulations. Given the multicultural and multilingual composition of ship crews, English serves as the lingua franca facilitating coordination among officers, crew members, port authorities, and other maritime stakeholders (Čulić-Viskota & Kalebota, 2013). In response to the necessity for standardized communication, the International Maritime Organization (IMO) developed the Standard Marine Communication Phrases (SMCP) to promote clarity and prevent misunderstandings in both routine operations and emergency situations (Limbong et al., 2024).

Proficiency in Maritime English, particularly oral communication, is critical for mitigating risks associated with miscommunication, which can result in navigational errors, operational inefficiencies, and maritime accidents (Yin et al., 2025). Nonetheless, many maritime students experience significant difficulties in attaining fluency and confidence. These challenges stem from limited exposure to authentic communicative contexts, pronunciation difficulties, complex technical vocabulary, and a lack of opportunities for real-time interactive practice (Celo et al., 2016). Conventional pedagogical methods—such as lectures, scripted dialogues, and role-plays—tend to prioritize linguistic accuracy over spontaneous communication, thereby inadequately developing learners' interactive competence (Halali et al., 2022).

Recent advancements in artificial intelligence (AI) have introduced innovative avenues for language learning enhancement. AI-driven conversational agents, such as OpenAI's ChatGPT, provide interactive and contextually adaptive environments in which learners can practice speaking, obtain immediate corrective feedback, and engage with domain-specific scenarios (Aljanabi, 2023). Empirical research within English for Specific Purposes (ESP) and English as a Foreign Language (EFL) context suggests that AI chatbots contribute positively to improvements in speaking fluency, vocabulary acquisition, and learner motivation, while also reducing communication anxiety (Pikhart, 2020).

Despite these promising outcomes, limited research has examined the use of ChatGPT specifically within Maritime English education. There remains a notable gap regarding the impact of AI-assisted interaction on students' speaking fluency, grammatical accuracy, and affective factors such as confidence and engagement, all of which are essential for communicative competence in maritime operations. This gap highlights the need for empirical studies focusing on AI-mediated oral practice in ESP contexts characterized by specialized terminology and realistic situational demands.

Therefore, changes in the field of education are a response to students where technology is aligned to create new opportunities creatively and innovatively. Technology can facilitate student interaction in the learning process and create an interactive and communicative classroom environment (Rojabi & Femilia, 2023). Learning at different times and places is characteristic of education in the Revolutionary Era (Marzuki et al., 2022). With the advancement of artificial intelligence (AI), ChatGPT presents a promising solution to enhance students' speaking skills in English learning (Ariati, 2024). ChatGPT, an AI-powered conversational tool, can simulate interactive dialogues, provide instant feedback, and create real-world maritime communication scenarios. By engaging in AI-driven conversations, students can improve their pronunciation, fluency, and confidence in using Maritime English in professional settings (Ratnaningsih et al., 2025).

So, this study aimed to investigate the integration of ChatGPT into Maritime English instruction to enhance students' speaking proficiency. Specifically, it examined the extent to which ChatGPT-assisted learning improves students' speaking fluency, grammatical accuracy, and

vocabulary use in Maritime English communication. In addition, the study explored students' perceptions of ChatGPT's role in fostering confidence, engagement, and motivation during speaking activities. Through these inquiries, the research sought to contribute to the growing body of literature on AI-assisted language learning and provide pedagogical insights for integrating emerging technologies into Maritime English education.

METHODS

This study employed Classroom Action Research (CAR) as its research design. CAR is a reflective and cyclical research approach aimed at improving teaching and learning practices through systematic planning, implementation, observation, and reflection in authentic classroom contexts (Manfra, 2019). In this study, CAR was used to investigate how the integration of ChatGPT into Maritime English instruction could enhance students' speaking proficiency. The research was conducted over four cycles, each consisting of the stages of planning, action, observation, and reflection.

The research was conducted during the 2024/2025 academic year at the Ship Operation Engineering Technology (Teknologi Rekayasa Operasional Kapal) study program of a maritime polytechnic in Indonesia. A total of 69 fourth-semester students (52 male and 17 female students, aged 19–21 years) participated in the study. All participants were enrolled in the Maritime English course and had previously completed introductory English for Specific Purposes (ESP) modules. The instruction spanned six 90-minute sessions, focusing on three thematic topics relevant to maritime communications such as safety equipment drills, the function of navigation equipment, and meteorological condition. Each topic was taught across two sessions.

The study followed four CAR cycles, as summarized in Table 1, to ensure iterative improvement and systematic observation.

TABLE 1 | Summary of the Classroom Action Research Cycles

Cycle	Focus and Activities	Data Collected
Cycle 1 - Planning	Identified students' speaking difficulties through a pretest and diagnostic questionnaire; designed the intervention by integrating ChatGPT conversation practice; prepared maritime-specific topics and rubrics for assessment.	Pretest scores, questionnaire data
Cycle 2 - Action	Implemented ChatGPT-assisted speaking practice. Students used ChatGPT to simulate real-life maritime dialogues based on assigned topics. The instructor facilitated discussions and provided feedback.	Observation notes, audio recordings

Cycle 3 - Observation	Observed student performance and interaction patterns using an observation checklist. Collected student feedback and lecturer evaluations using validated rubrics for fluency, accuracy, vocabulary, and pronunciation (Fadila & Trisno, 2025).	Observation data, rubric scores, field notes
Cycle 4 – Reflection and Evaluation	Analyzed and reflected on the results of the intervention. Conducted a post-test and follow-up questionnaire to evaluate improvement and perceptions. Discussed findings with the course lecturer.	Post-test scores, reflection data

The study employed several instruments to collect quantitative and qualitative data. A pretest and post-test were administered to measure students' speaking performance before and after the ChatGPT-assisted intervention. In each test, students performed a short maritime communication task evaluated using a validated speaking rubric comprising four criteria—fluency, grammatical accuracy, vocabulary, and pronunciation—rated on a 1–5 scale (adapted from [Fadila & Trisno, 2025](#)). The rubric demonstrated high reliability, with a Cronbach's alpha value of 0.87, and inter-rater reliability was confirmed through Cohen's kappa ($\kappa = 0.82$) after two independent raters assessed each performance. Additionally, two sets of questionnaires (administered pre- and post-intervention) were used to gather information on students' perceptions, speaking difficulties, and attitudes toward ChatGPT use. The questionnaire items were rated on a 5-point Likert scale, and their content validity was ensured through expert review by three Maritime English lecturers. Finally, an observation sheet was utilized to record classroom interactions, engagement levels, and language use during ChatGPT-based activities, with observations conducted by the researcher and validated through peer debriefing with the course lecturer.

Data collection was conducted over six consecutive sessions spanning three weeks. In the first week, a pretest and diagnostic questionnaire were administered to identify students' initial speaking proficiency and perceptions. During the second and third weeks, ChatGPT-assisted learning activities were implemented as the core intervention. At the end of Week 3, a post-test, final questionnaire, and reflection session were carried out to evaluate learning outcomes and gather feedback on the use of ChatGPT in Maritime English instruction. All classroom sessions were audio-recorded to ensure data accuracy, and field notes were maintained throughout the process to support the qualitative analysis.

Quantitative data from the pretests and post-tests were analyzed using SPSS version 22. A paired samples t-test was conducted to determine whether there was a statistically significant difference in students' speaking performance

before and after the ChatGPT-assisted intervention. Descriptive statistics, including mean, standard deviation, and percentage, were used to analyze the questionnaire responses. Meanwhile, qualitative data obtained from classroom observations and open-ended questionnaire items were analyzed thematically to identify recurring patterns related to student engagement, confidence, and perceived learning improvement.

To ensure the credibility and trustworthiness of the findings, several validation strategies were applied. Triangulation was used by combining quantitative data (test scores) and qualitative data (observations and questionnaires). Peer debriefing was conducted regularly, allowing the researcher to discuss observations and interpretations with a peer lecturer to ensure consistency. Member checking was also implemented, in which selected students reviewed summaries of observation notes to confirm their accuracy. Furthermore, all research instruments were validated through expert review to ensure they measured the intended constructs effectively.

Ethical approval for this study was granted by the institution's Research Ethics Committee. Prior to participation, all students were informed about the study's objectives and procedures, and informed consent was obtained. Participants were assured of anonymity and confidentiality, and their participation was entirely voluntary, with the option to withdraw at any time without any academic penalty.

RESULTS AND DISCUSSION

Cycle 1 – Problem Identification

A pre-test was conducted to identify students' challenges in speaking Maritime English before implementing any instructional interventions. This initial assessment aimed to evaluate their baseline performance across four key aspects of speaking skills—grammar, vocabulary, fluency, and comprehension—within the context of maritime-related topics such as Safety Equipment, Navigation Equipment, and Meteorological Conditions. The results of the pre-test provided valuable insights into the areas where students struggled most, serving as a foundation for designing targeted learning strategies to enhance their communicative competence in Maritime English.

The test was about the students' performance in explaining and performing dialogs with their peers about the maritime-related topics, they were safety equipment, navigation equipment, and meteorological condition. The aspects in the speaking test were the grammar, vocabulary, fluency and comprehension. The sample of dialogs were transcribed and analyzed as the qualitative data. Below was one of the samples of dialog transcriptions which talked about the safety equipment on board.

Transcription data 1

Student 6: *Eh... good morning, my friend. Today we talk about safety, yes?*

Student 7: *Yes, about safety equipment on the sheep... eh, ship.*

Student 6: (smiles) Yes, ship. Okay. The... uh... lifeboat is use for... for escape when the ship is sinking.
 Student 7: Yes, and... the lifejacket, we must wear it when... when emergency coming.
 Student 6: Coming... or happen?
 Student 7: Ah, yes, when emergency happen.
 Student 6: Okay. And fire extinguisher... it use to stop fire.
 Student 7: Yes, but sometimes we no know how to use correctly.
 Student 6: True. The officer must train us how to operate it.
 Student 7: Yes, sometimes I forget the name of equipment, because many... many name.

Based on the observation in transcription data 1, the students demonstrated several speaking difficulties commonly found among English language learners in a maritime context. Grammatical errors were evident in utterances such as “It use for escape” instead of “It is used for

escaping,” showing limited mastery of verb forms and passive constructions. Pronunciation problems appeared in the confusion between similar sounds, as in “sheep” for “ship,” which affected intelligibility. Vocabulary limitations were also noticeable when students hesitated or used inaccurate expressions like “when emergency coming.” In terms of fluency, frequent pauses and fillers (e.g., “uh,” “for... for escape”) indicated lack of confidence and insufficient automaticity in speech production. Despite these weaknesses, students showed awareness of self-correction, as seen when one learner replaced “coming” with “happen.” Overall, the observation revealed that students’ speaking difficulties mainly stemmed from limited grammatical accuracy, restricted vocabulary range, pronunciation interference, and low fluency, which hindered effective oral communication during the activity.

Then, based on the dialogs that performed by the students, the pretest scores were taken. Then, the pretest result was shown in [table 2](#).

TABLE 2 | Pretest Result

Topic	Grammar	Vocabulary	Fluency	Comprehension	Overall Avg
Safety Equipment	70.2	75.6	73.8	77.4	74.3
Navigation Equipment	73.5	78.9	75.2	80.1	76.9
Meteorological Conditions	69.8	74.4	72.1	76.3	73.2

The results presented in [table 2](#) showed the results of the speaking performance assessment across three Maritime English topics—Safety Equipment, Navigation Equipment, and Meteorological Conditions—showed varying levels of improvement in different aspects of language use. For the topic of Safety Equipment, students achieved average scores of 70.2 in grammar, 75.6 in vocabulary, 73.8 in fluency, and 77.4 in comprehension, with an overall mean score of 74.3. In Navigation Equipment, performance slightly improved across all areas, with scores of 73.5 in grammar, 78.9 in vocabulary, 75.2 in fluency, and 80.1 in comprehension, resulting in the highest overall average of 76.9. Meanwhile, for Meteorological Conditions, students obtained scores of 69.8 in grammar, 74.4 in vocabulary, 72.1 in fluency, and 76.3 in comprehension, with an overall average of 73.2. These results suggested that students demonstrated stronger performance in comprehension and vocabulary, particularly in topics involving navigation-related communication.

Then, a questionnaire was administered to identify students’ challenges in speaking Maritime English. There were five items in the questionnaire that investigated the students’ speaking difficulties. The first was to determine the limited vocabulary that the students had, with the statement, “I often struggled to use specific maritime vocabulary when speaking English.” The second concerned the students’ self-confidence when speaking English, expressed in the statement, “I felt nervous or lacked confidence when I had to speak English in front of others.” The third addressed grammatical construction difficulties through the statement, “I found it difficult to create grammatically correct sentences when speaking English.” The fourth referred to the inability to understand maritime communication contexts, as reflected in the statement, “I had trouble understanding the context or meaning of maritime communication expressions.” The last

item focused on the fear of making mistakes, represented by the statement, “I was afraid of making mistakes when speaking English, which made me hesitate to speak.” Each of these questionnaire items was designed using a 5-point Likert scale (1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly Agree).

TABLE 3 | Questionnaire Result Cycle 1

Number	Speaking Difficulty	Average Rating (out of 5)
1	Limited maritime vocabulary	4.8
2	Lack of confidence in speaking English	4.1
3	Difficulty in constructing grammatically correct sentence	4
4	Inability to understand maritime communication contexts	3
5	Fear of making mistakes during conversations	4

Based on the data presented in [Table 3](#): Questionnaire Result Cycle 1, The results presented in Table 3 showed the students perceived speaking difficulties during Cycle 1 of the study. The highest-rated difficulty was limited maritime vocabulary, with an average rating of 4.8 out of 5, indicating that most students struggled significantly with recalling and using technical maritime terms in communication. The next major challenges were a lack of confidence in speaking English (4.1) and difficulty in constructing grammatically correct sentences (4.0), suggesting that linguistic competence

and self-assurance were closely linked issues affecting their speaking performance. The fear of making mistakes during conversations also received a relatively high rating (4.0), reflecting students' anxiety and hesitation in oral communication. In contrast, inability to understand maritime communication contexts had the lowest rating (3.0), implying that while students generally understood the situational aspects of maritime exchanges, their main barriers lay in vocabulary, grammar, and confidence. Overall, the data indicate that students' speaking difficulties were primarily linguistic and psychological rather than contextual.

In conclusion, the findings from both the pretest and questionnaire highlighted that the students' primary speaking difficulties before the intervention were linguistic and affective rather than contextual. These baseline results justified the need for targeted instruction that aimed to enhance maritime vocabulary, grammatical accuracy, and communicative confidence through technology-assisted learning.

Cycle 2 – Implementation of ChatGPT – Assisted Activities

In this phase, students had the ChatGPT integration in the speaking practices. They simulated the dialogs based on three maritime English topics. They are the safety equipment, navigation equipment and meteorological conditions. The questionnaires items were ChatGPT was helpful in supporting my speaking practice in maritime English, Using ChatGPT helped me learn and use more specific maritime vocabulary, I felt more confident speaking English after practicing with ChatGPT, ChatGPT helped me construct more grammatically correct and natural-sounding sentences, and the speaking scenarios provided by ChatGPT were relevant to maritime communication situations. Each of those items designed for a 5-point Likert scale (1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly Agree).

TABLE 4 | Questionnaire Result Cycle 2

Number	Aspect	Average Rating (out of 5)
1	Helpfulness of ChatGPT	4.6
2	Vocabulary improvement	4.4
3	Confidence in speaking	4.2
4	Grammar assistance	4.3
5	Relevance of maritime scenarios	4.5

The data in [table 4](#) showed the students' experience using ChatGPT in maritime English learning was evaluated using a 5-point Likert scale, and the results showed a generally high level of satisfaction across all aspects. The results of the post-intervention questionnaire indicate that students responded positively to the integration of ChatGPT in Maritime English learning. As shown in the data, the helpfulness of ChatGPT received the highest average rating of 4.6, suggesting that students perceived the tool as highly effective in supporting their learning process. The relevance of maritime scenarios also scored high at 4.5, reflecting that the AI-generated dialogues and contexts were appropriate and aligned with

real-life maritime communication situations.

In terms of language development, students reported notable improvement in vocabulary (4.4) and grammar assistance (4.3), showing that ChatGPT effectively supported their linguistic competence by providing meaningful exposure to technical terms and grammatical structures. Additionally, the confidence in speaking aspect received a rating of 4.2, indicating that interaction with ChatGPT helped reduce students' hesitation and anxiety during oral communication, although some learners may still require further practice to build stronger self-assurance. The transcription data 2 below was from the observation activities in cycle 2.

Transcription data 2

- Student 6: *Today we are talking about safety equipment on board. I think the lifejacket is the most important one because it help—eh, helps us float in the water.*
- Student 7: *Yes, and also we must know how to wear it correctly. Some cadets still not know—uh, don't know how to adjust the straps.*
- Student 6: *Right. And the lifebuoy is used when someone fall overboard. We can throw it quickly to the sea.*
- Student 7: *Yes, it has line for pulling the person back to the ship. Before, I didn't know the word line for rope, but ChatGPT teach me that.*
- Student 6: *Me too. I also learn the word buoyancy and fire extinguisher. The extinguisher is used for stop—uh, stopping the fire in emergency.*
- Student 7: *Yes, and now I feel more confident to speak because before I always afraid to make mistake.*

The observation notes indicated clear evidence of students' language development in several key areas of speaking performance. First, there was notable improvement in vocabulary use, as students began to employ more accurate and topic-specific maritime terms such as *buoyancy, lifebuoy, and fire extinguisher*.

This suggests that consistent exposure to technical language during guided practice enhanced their lexical repertoire and helped them communicate more precisely about shipboard safety. Second, grammar improvement was evident through self-correction behaviors, for instance when students adjusted phrases like “*it help*” to “*it helps*” and “*stop*” to “*stopping*.” These self-repairs show growing grammatical awareness and internalization of feedback provided during instruction.

Furthermore, the observation reflected increased confidence and fluency in oral communication. Students participated actively, expressed ideas more spontaneously, and demonstrated reduced hesitation, indicating progress in overcoming speaking anxiety. Although minor issues such as verb tense inconsistency and omission of articles persisted, these did not significantly hinder comprehension. Overall, the analysis confirmed that the students made steady progress in vocabulary acquisition, grammatical accuracy, and speaking confidence, reflecting the effectiveness of communicative practice and interaction-based learning in improving their Maritime English proficiency.

In general, these findings in this cycle demonstrated that ChatGPT contributed positively to both the linguistic and affective aspects of Maritime English learning. The tool was perceived as a practical and engaging aid that enhanced vocabulary mastery, grammatical accuracy, and speaking confidence through authentic, scenario-based communication tasks.

Cycle 3 – Improvement in Performance and Learner Engagement

In Cycle 3, the focus shifted toward evaluating the impact of ChatGPT-integrated speaking practices through performance observation and feedback. During this phase, the lecturer conducted systematic observations by video-recording the students' speaking performances in a controlled classroom setting.

Students were assigned to work individually and in pairs to simulate maritime-specific dialogues using ChatGPT as a practice partner prior to the recorded assessment. The ChatGPT provided them with interactive prompts, vocabulary suggestions, and corrective feedback, helping them prepare for real-time verbal communication in a maritime context.

The speaking tasks in this cycle were organized around three essential Maritime English topics that are fundamental for effective communication at sea. The first topic, Safety Equipment, required students to describe and explain the function, use, and location of life-saving appliances such as lifejackets, fire extinguishers, and lifeboats. The second topic, navigation equipment, focused on technical dialogues involving the use of radar, GPS (Global Positioning System), AIS (Automatic Identification System), ECDIS (Electronic Chart Display and Information System), and the compass on the bridge. Last, the third topic, meteorological conditions, engaged students in reporting and interpreting weather updates by using appropriate maritime terminology related to wind direction, sea state, fog, storms, and visibility. Together, these topics aimed to enhance students' mastery of technical vocabulary and situational communication skills relevant to shipboard operations.

During the assessment, students were asked to present short dialogues reflecting realistic on-board scenarios. Their performances were recorded to allow for detailed analysis across four criteria: grammar, vocabulary, fluency, and comprehension. The lecturer then analyzed the recordings, using a standardized rubric to assess progress. Additionally, peer and lecturer feedback sessions were conducted after the recordings. Students were shown clips of their performances and asked to reflect on their speaking—what went well, where they hesitated, and how they could improve. The observation sample activities were shown in transcription data 3 for the students' dialogs and transcription data 4 for lecturer feedback and discussion session.

Transcription data 3

Student 6: *Good morning. Have you checked the safety equipment in the engine room today?*

Student 7: *Yes, I already check—uh, checked—the fire extinguishers and the emergency exit light. Everything is in good condition.*

Student 6: *Good. Don't forget to inspect the lifejackets in the locker. Some of them maybe—uh, may be—need to replace because the straps are broken.*

Student 7: *Oh yes, I saw one lifejacket with torn belt. I will report it to the duty officer after this.*

Student 6: *Also, the lifebuoy on the aft deck must be checked every week. Last week it was missing the line.*

Student 7: *Yes, sir. I will make sure it's complete now. The lifebuoy is important when man overboard happen—uh, happens.*

Student 6: *Correct. We must always be ready for any emergency.*

Transcription data 4

Lecturer: *Very good, cadets. You both demonstrated clear understanding of the topic. I can see improvement in your vocabulary, especially with technical words like lifebuoy, lifejacket, and emergency exit light.*

Student 7: *Thank you, mam. I try to remember the correct terms from our last practice with ChatGPT.*

Lecturer: *That's great! However, be careful with verb forms. For example, you said "I already check" — it should be "I have already checked."*

Student 6: *Yes, mam. I will remember that next time.*

Lecturer: *Also, good effort on fluency. There were only small pauses when you tried to find the right word, which is normal. How did you feel during the recording?*

Student 7: *At first, I feel nervous, but after speaking a few minutes, I become more confident.*

Lecturer: *Excellent. That's progress. You are speaking more naturally now, and your comprehension is improving too — you understood each other without repeating instructions. Keep practicing to strengthen grammar and fluency.*

The dialogue and feedback session in transcription data 3 and 4 above revealed clear evidence of students' language development and communicative progress in the topic of Safety Equipment on Board. Both students demonstrated an improved command of maritime-specific vocabulary, as they used technical terms such as lifebuoy, lifejacket, fire extinguisher, and emergency exit light. This indicated that the learners had become more familiar with shipboard terminology through repeated exposure and practice. Their comprehension was also strong, as shown by the smooth exchange of information and their ability to respond appropriately without needing clarification or repetition.

In terms of grammar, minor errors were observed in verb forms and modal constructions, for instance, "I already check" instead of "I have already checked" and "when man overboard happen" instead of "happens." However, the students' self-corrections during speech (e.g., "check—uh, checked") demonstrated growing grammatical awareness and the ability to monitor their own language use in real time. This self-correcting behavior reflected meaningful improvement in their linguistic competence.

Regarding fluency, both students maintained a natural conversational flow with minimal hesitation, showing better confidence and rhythm compared to earlier practice sessions. Although there were slight pauses when they searched for the

correct word, these did not disrupt communication. The students also showed increased confidence, as noted when Ahmad mentioned that he felt *“more confident”* after the recording. The lecturer’s feedback confirmed these observations, emphasizing progress in vocabulary accuracy,

fluency, and comprehension while acknowledging that grammar still required consistent practice.

Then, the result summary of the speaking tasks with feedback and discussion was shown in [table 5](#).

TABLE 5 | Summary of Speaking Tasks Average Scores

Topic	Grammar	Vocabulary	Fluency	Comprehension	Overall Avg
Safety Equipment	79.4	84.2	82.1	86.5	83.0
Navigation Equipment	82.7	88.5	85.3	90.2	86.7
Meteorological Conditions	78.6	83.1	80.8	85.0	81.9

The results presented in [table 5](#) showed a clear improvement in students’ speaking performance across all Maritime English topics after the ChatGPT-assisted learning intervention. Among the assessed aspects—grammar, vocabulary, fluency, and comprehension—all showed notable gains compared to the pretest results, indicating the positive impact of AI-supported instruction on students’ speaking proficiency.

The Navigation Equipment topic achieved the highest overall average score of 86.7, with particularly strong results in comprehension (90.2) and vocabulary (88.5). This suggests that students became more confident and accurate when communicating in familiar, technical maritime contexts. Similarly, performance in Safety Equipment was strong, with an overall mean of 83.0, reflecting consistent improvement in all areas, especially in comprehension (86.5) and vocabulary (84.2).

Meanwhile, the Meteorological Conditions topic, though slightly lower with an overall average of 81.9, still demonstrated substantial progress, indicating better understanding and use of meteorological terms and expressions.

Overall, the data indicate that students made the most progress in comprehension and vocabulary, suggesting that ChatGPT effectively enhanced their understanding of maritime contexts and exposure to specialized terminology. Improvements in grammar and fluency further reveal growing linguistic competence and confidence in oral communication. These findings confirm that ChatGPT-assisted activities significantly supported students’ development of both language accuracy and communicative performance in Maritime English.

In this cycle, not only observation from the lecturer, but also the questionnaire which distributed to the students was also the instrument to measure the performance and feedback. The questionnaire items were increase in participation rate with the statement *“I became more active in speaking activities”*, then the improving used of maritime terms in the statement *“Now, I am able to use maritime specific terms vocabulary when speaking”*. Next items were fluency improvement, better grammar and coherence in response, and speaking development. The statements were *“I feel that my spoken English has become more fluent during maritime English practice”*, *“my sentence structure has improved when I speak English in a maritime context”*, and *“practicing speaking over several sessions helped me express ideas more clearly in English.”* The summary result was shown in table

6.

TABLE 6 | Questionnaire Result Cycle 3

Number	Aspect	Average Rating (out of 5)
1	Increase in participation rate	4.8
2	Improved used in maritime terms	4.5
3	Fluency improvement	4.4
4	Better grammar and coherence in response	4.5
5	Speaking development	4.7

Based on the data presented in [Table 6](#), the results of the questionnaire in Cycle 3 reflected improved engagement, with the highest ratings in participation (M = 4.8) and speaking development (M = 4.7). This suggests that the integration of ChatGPT fostered greater learner engagement and self-efficacy, supporting prior findings on AI-enhanced motivation (Wang & Kasper, 2022). Importantly, fluency (M = 4.4) and grammatical coherence (M = 4.5) also improved, indicating that AI-based corrective feedback and repetition cycles contributed to automaticity in speech production (DeKeyser, 2007). However, data in cycle 3 revealed that some students experienced hesitation when shifting from AI-supported dialogues to peer interactions, underscoring the need to combine AI practice with collaborative speaking tasks to enhance authentic communicative readiness.

Cycle 4 – Evaluation (Reflections and Performance Analysis)

Finally, post-test was conducted to identify students’ challenges in speaking Maritime English before implementing any instructional interventions. This initial assessment aimed to evaluate their baseline performance across four key aspects of speaking skills such as grammar, vocabulary, fluency, and comprehension within the context of maritime-related topics such as Safety Equipment, Navigation Equipment, and Meteorological Conditions.

The post-test was about the students’ performance in explaining and performing dialogs with their peers about the maritime-related topics, they were safety equipment, navigation equipment, and meteorological condition. The observation sample activity was shown in transcription data 5 for the students’ post-test.

Transcription data 5

- Student 6: *Good morning. Today we need to talk about the safety equipment on board, right?*
- Student 7: *Yes, right. I will start first. Safety equipment are very important to keep the crew safe in emergency situation.*
- Student 6: *Yes, I agree. For example, we have lifejackets, lifebuoy, and also fire extinguisher. All must be checked regularly by the crew.*
- Student 7: *(nodding) Yes, the officer in charge should make sure everything is in good condition. The lifejackets must be, uh, placed in the locker near the muster station.*
- Student 6: *Yes, and lifebuoy should be ready on the deck for man overboard accident.*
- Student 7: *We also use fire extinguisher to fight the fire in engine room. But some cadets, they not remember to check the pressure gauge.*
- Student 6: *(laughs) Yes, that's true. Sometimes we forget small things like that. During my duty, I always check the label and expiry date.*
- Student 7: *Good practice. Also, the emergency light and alarm must be test every week.*
- Student 6: *Oh yes, test every week. And we should record it in the log book or report to the engineer officer.*
- Student 7: *So, to conclude, safety equipment must be ready anytime, and every crew have responsibility to maintain it.*
- Student 6: *Yes, safety first before everything. Thank you, Ahmad.*
- Student 7: *Thank you, Rafi.*

Based on the post-test speaking performance, the students demonstrated an overall improvement in their ability to explain and perform dialogues related to maritime topics, particularly safety equipment. In terms of grammar, their sentences were mostly accurate, though minor mistakes were observed in subject-verb agreement and article usage, such as “*safety equipment are*” instead of “*safety equipment is.*” The vocabulary aspect showed that students appropriately used maritime-related terms like fire extinguisher, pressure gauge, and locker, indicating topic familiarity; however, some repetition suggested a limited lexical range. Regarding fluency, the students spoke smoothly with only a few pauses and filler words such as “*uh*” and “*yes, right,*” showing growing confidence in oral communication. Finally, their comprehension was strong, as both speakers responded logically and maintained coherence throughout the dialogue, despite slight delays in turn-taking. Overall, the students’ post-test performance reflected satisfactory progress across all speaking aspects, with only minor issues in grammatical precision and vocabulary variety. The aspects in the speaking test were the grammar, vocabulary, fluency and comprehension. The pretest result was shown in [table 7](#).

The post-test results, as shown in [Table 7](#), demonstrated a significant improvement in students’ speaking performance in Maritime English following the learning intervention. Each assessed component—grammar, vocabulary, fluency, and comprehension—shows measurable gains across all three maritime-related topics, indicating enhanced communicative competence and confidence in using English within maritime contexts.

TABLE 7 | Post-test Result

Topic	Grammar	Vocabulary	Fluency	Comprehension	Overall Avg
Safety Equipment	83.5	87.6	85.9	89.4	86.6
Navigation Equipment	86.2	91.3	88.7	93.5	88.9
Meteorological Conditions	82.4	86.9	84.6	88.2	85.5

For the Safety Equipment topic, students achieved an overall average of 86.6, a noticeable increase from their pre-test average of 74.3 and even higher than the initial post-test score of 83.0. Comprehension remained the strongest aspect (89.4), suggesting better understanding and response accuracy during oral tasks. Grammar and fluency also showed strong improvements, reaching 83.5 and 85.9, respectively. In the Navigation Equipment topic, students recorded the highest overall average of 89.9. This represents outstanding development, particularly in comprehension (93.5) and vocabulary (91.3), reflecting their growing familiarity with technical terminology and ability to use it appropriately in context. Grammar (86.2) and fluency (88.7) scores also indicate more accurate and natural language production during speaking tasks. The topic of Meteorological Conditions, which had the lowest scores in the pre-test, also saw substantial growth, with an overall average of 85.5. Grammar (82.4), vocabulary (86.9), and fluency (84.6) scores reveal increased confidence in discussing weather-related maritime topics, while comprehension (88.2) shows improved listening and interpretive skills during exchanges.

Overall, these post-test results reflect the effectiveness of the instructional strategies implemented to improve students’ Maritime English-speaking skills. The consistent upward trend across all language components and topics suggests that learners not only acquired new vocabulary and structures but also developed better fluency, comprehension, and communicative strategies suitable for maritime communication scenarios.

Then, the last questionnaire was distributed to the students. The items were related to the speaking performance improvement such as the improvement of the grammar accuracy, wider range of vocabulary in maritime context, higher comprehension and accurate responses to the questions and instruction in maritime English, and increased fluency with less pauses and hesitation.

The results of the Cycle 4 questionnaire indicated that students perceived substantial improvements in all key aspects of their Maritime English-speaking skills after the ChatGPT-assisted learning intervention. The highest average rating of 4.8 for grammar accuracy improvement suggested that learners acknowledged the effectiveness of the

instructional approach and AI feedback in enhancing their grammatical precision during oral communication. The aspect of wider maritime vocabulary also received a high rating of 4.6, showing that the integration of authentic maritime contexts and repeated AI-guided practice helped students acquire and apply domain-specific terminology more confidently.

TABLE 8 | Questionnaire Result Cycle 4

Number	Aspect	Average Rating (out of 5)
1	Grammar accuracy improvement	4.8
2	Wider maritime vocabulary	4.6
3	Higher comprehension and more accurate responses	4.7
4	Increasing of fluency	4.7

Meanwhile, higher comprehension and more accurate responses as well as increasing fluency both achieved an average rating of 4.7, indicating that students experienced clearer understanding of instructions, better message

interpretation, and smoother speech production. Overall, the consistently high scores across all aspects demonstrated that students not only improved linguistically but also gained greater communicative confidence and motivation. These outcomes confirmed that the ChatGPT-integrated instruction successfully supported both the linguistic and affective domains of Maritime English learning, aligning with previous findings that technology-enhanced and interactive learning environments fostered accuracy, fluency, and learner engagement.

Furthermore, to evaluate the effectiveness of the instructional intervention on students' English-speaking skills, a paired sample t-test was conducted comparing their performance before and after the treatment. The analysis focused on five critical aspects of speaking: grammar, vocabulary, fluency, comprehension, and the overall average score. By comparing the pretest and post-test results, the statistical test aimed to determine whether the observed improvements were significant and consistent across all measured components. The statistical test result was presented in [table 9](#).

TABLE 9 | Result of paired sample t-test

Aspects	Pretest Mean	Posttest Mean	Mean Difference	Sig (2 -tailed)	Interpretation
Grammar	71.17	84.03	+12.86	< 0.05	Significant improvement
Vocabulary	76.30	88.60	+12.30	< 0.05	Significant improvement
Fluency	73.70	86.40	+12.70	< 0.05	Significant improvement
Comprehension	77.93	90.37	+12.44	< 0.05	Significant improvement
Overall Average	74.80	87.00	+12.20	< 0.05	Significant overall improvement

The results of the Cycle 4 questionnaire indicated that students perceived substantial improvements in all key aspects of their Maritime English-speaking skills after the ChatGPT-assisted learning intervention. The highest average rating of 4.8 for grammar accuracy improvement suggested that learners acknowledged the effectiveness of the instructional approach and AI feedback in enhancing their grammatical precision during oral communication. The aspect of wider maritime vocabulary also received a high rating of 4.6, showing that the integration of authentic maritime contexts and repeated AI-guided practice helped students acquire and apply domain-specific terminology more confidently. Meanwhile, higher comprehension and more accurate responses as well as increasing fluency both achieved an average rating of 4.7, indicating that students experienced clearer understanding of instructions, better message interpretation, and smoother speech production. Overall, the consistently high scores across all aspects demonstrated that students not only improved linguistically but also gained greater communicative confidence and

motivation. These outcomes confirmed that the ChatGPT-integrated instruction successfully supported both the linguistic and affective domains of Maritime English learning, aligning with previous findings that technology-enhanced and interactive learning environments fostered accuracy, fluency, and learner engagement.

The findings across the four research cycles collectively demonstrated a steady and significant improvement in students' Maritime English-speaking proficiency, confirming the effectiveness of ChatGPT-assisted learning.

In Cycle 1, both pretest and questionnaire results revealed that students' main speaking difficulties had been linguistic (limited vocabulary and grammar accuracy) and affective (lack of confidence and anxiety), rather than contextual or situational. These baseline findings justified the need for a targeted instructional approach that focused on vocabulary enrichment, grammatical precision, and communicative confidence. According to [Thi Thuy Ha & Huu Duc \(2024\)](#), speaking difficulties often arose from inadequate linguistic resources and affective barriers, which could be mitigated

through interactive and feedback-rich environments. Therefore, the initial diagnosis set the foundation for employing AI-assisted tasks to provide individualized practice and immediate feedback—factors known to enhance language acquisition (Abusahyon et al., 2023).

By Cycle 2, noticeable progress was evident in vocabulary acquisition, grammatical accuracy, and speaking confidence. This improvement reflected the effectiveness of communicative practice and interaction-based learning integrated with ChatGPT. The results aligned with Bau & Hien (2023), which emphasized that learning occurred through meaningful interaction and scaffolding. ChatGPT served as a dynamic conversational partner, facilitating such scaffolding by providing authentic maritime communication scenarios. Similar findings were reported by Nugroho et al. (2023), who found that AI tools promoted learner engagement and linguistic development in English for Specific Purposes (ESP) contexts through authentic task simulation.

During Cycle 3, the gains became more pronounced across all assessed aspects—grammar, vocabulary, fluency, and comprehension. Students' recorded speaking performances and observational data revealed more accurate sentence construction, wider lexical use, smoother speech delivery, and stronger understanding of maritime discourse. These improvements demonstrated that iterative AI-mediated interactions had fostered automaticity in language use, consistent with Eragamreddy (2025) theory that fluency developed through repeated, meaningful exposure so the integration of ChatGPT simulated this interactive negotiation process, enabling students to self-correct and refine their linguistic output.

In Cycle 4, both quantitative and qualitative data confirmed substantial progress across all language components. The post-test results, supported by the paired sample t-test, revealed statistically significant improvements ($p < 0.05$) in grammar (+12.86), vocabulary (+12.30), fluency (+12.70), and comprehension (+12.44), leading to an overall mean increase from 74.80 to 87.00. This upward trend confirmed that the instructional intervention had produced a systematic and cumulative impact on learners' speaking performance. The highest gain in grammar accuracy suggested that explicit corrective feedback—provided through ChatGPT's responses and lecturer reflections—had helped students internalize correct linguistic forms. This aligned with Van Beuningen (2011) hypothesis and Abdelhalim & Alsehibany (2025), emphasizing the role of comprehensible input and pushed output in developing grammatical competence.

The questionnaire results in Cycle 4 further validated these findings from an affective perspective. Students reported notable increases in confidence, fluency, and comprehension, with average ratings above 4.6 out of 5 across all aspects. This shift indicated that ChatGPT not only enhanced linguistic skills but also reduced anxiety and improved motivation—factors central to affective learning theories (Hayashi & Sato, 2024). The tool's nonjudgmental and interactive nature provided a safe space for practice, thereby fostering willingness to communicate—a critical predictor of speaking success (Zhang et al., 2024).

Overall, these findings substantiated the pedagogical value of AI integration in ESP instruction, particularly within the Maritime English domain. The consistent improvements observed from Cycle 1 to Cycle 4 highlighted the synergy between technology-enhanced interaction, communicative task design, and reflective feedback. Similar studies by Cavazos et al. (2024) and (Almogren et al., 2024) demonstrated that ChatGPT could serve as a transformative medium for language learning, enabling contextualized, self-paced, and confidence-building practice.

In summary, the integration of quantitative and qualitative findings confirmed that ChatGPT-assisted instruction had effectively addressed the linguistic and affective challenges identified at the start of the research. Through iterative practice, personalized feedback, and authentic maritime communication tasks, learners developed measurable gains in grammar, vocabulary, fluency, and comprehension. These results reinforced the theoretical assertion that AI-supported communicative learning enhanced both the cognitive and emotional dimensions of second language acquisition, providing a sustainable model for Maritime English education and broader ESP contexts.

CONCLUSION

The use of ChatGPT in Maritime English-speaking classes has shown to be an effective tool for enhancing vocabulary, grammar, fluency, and learner confidence. The structured CAR methodology enabled targeted improvements through iterative cycles of action, observation, and reflection. Future implementations should consider expanding the use of AI-based simulations across more maritime communication scenarios and integrating peer-collaborative tasks using ChatGPT.

ACKNOWLEDGEMENTS

The authors express their gratitude to all of the students who participated in this study as well as to a few colleagues for their insightful remarks, recommendations, and significant revisions made to this manuscript. Lastly, we would like to thank the anonymous reviewers of this article for their informative remarks and notes.

REFERENCES

- Abdelhalim, S. M., & Alsehibany, R. (2025). Integrating ChatGPT for vocabulary learning and retention: A classroom-based study of Saudi EFL learners. *Language Learning & Technology*, 29(1), 1–24. <https://doi.org/10.64152/10125/73635>
- Abusahyon, A. S. E., Alzyoud, A., Alshorman, O., & Al-Absi, B. (2023). AI-driven Technology and Chatbots as Tools for Enhancing English Language Learning in the Context of Second Language Acquisition: A Review Study. *International Journal of Membrane Science and Technology*, 10(1), 1209–1223. <https://doi.org/10.15379/ijmst.v10i1.2829>

- Aljanabi, M. (2023). ChatGPT: Future Directions and Open possibilities. *Mesopotamian Journal of CyberSecurity*, 2023(January), 16–17. <https://doi.org/10.58496/MJCS/2023/003>
- Almogren, A. S., Al-rahmi, W. M., & Ahmed, N. (2024). Heliyon Exploring factors influencing the acceptance of ChatGPT in higher education: A smart education perspective. *Heliyon*, 10(11), e31887. <https://doi.org/10.1016/j.heliyon.2024.e31887>
- Ariati, J. (2024). Integrating ChatGPT into English Language Teaching: a Comprehensive Study on Its Impact, Perceptions, and Ethical Implications. *Proceedings of The Fourth International Conference on English Language Education*, August, 132–138. <https://proceedings.uin-alauddin.ac.id/index.php/iconole/Iconole04/paper/view/1573>
- Bau & Hien. (2023). Exploring the Efficacy of ChatGPT in Language Teaching. 14(2), 156–167. <https://doi.org/10.54855/acoj.2314210>
- Cavazos, J. T., Hauck, K. A., Baskin, H. M., & Bain, C. M. (2024). ChatGPT Goes to College: Exploring Student Perspectives on Artificial Intelligence in the Classroom. *Teaching of Psychology*, August. <https://doi.org/10.1177/00986283241268829>
- Celo, E., Danaj, L., & Celo, J. (2016). Maritime English - Teaching Competencies in Facilitating Difficulties with Students. *European Journal of Multidisciplinary Studies*, 1(2), 95. <https://doi.org/10.26417/ejms.v1i2.p95-103>
- Čulić-Viskota, A., & Kalebota, S. (2013). Maritime English – What Does It Communicate? *Transactions on Maritime Science*, 2(2), 109–114. <https://doi.org/10.7225/toms.v02.n02.003>
- Eragamreddy, D. N. (2025). Interactive AI-Driven Storytelling for Language Development. *International Journal of Social Science Humanity & Management Research*, 04(02), 313–333. <https://doi.org/10.58806/ijsshmr.2025.v4i2n08>
- 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>
- Halali, A. A. S., Ismail, L., Samad, A. A., Razali, A. B., & Noordin, N. (2022). Challenges in Academic Speaking for Non-Native Speakers: The Case of Libyan Students Studying in Malaysia. *Pertanika Journal of Social Sciences and Humanities*, 30(December), 43–62. <https://doi.org/10.47836/PJSSH.30.S1.03>
- Hayashi, K., & Sato, T. (2024). The Effectiveness of ChatGPT in Enhancing English Language Proficiency and Reducing Second Language Anxiety (L2). *WorldCALL2023: Official Conference Proceedings, March*, 201–208. <https://doi.org/10.22492/issn.2759-1182.2023.23>
- Limbong, S., Chandra, A., Tandibura, F., & Hafita, Y. A. (2024). The Convergence of Maritime English and Green Shipping: Enhancing Collaboration Through Effective Communication (Issue ICMaD). *Atlantis Press International BV*. https://doi.org/10.2991/978-94-6463-628-4_22
- Manfra, M. M. G. (2019). Action Research and Systematic, Intentional Change in Teaching Practice. *Review of Research in Education*, 43(1), 163–196. <https://doi.org/10.3102/0091732X18821132>
- Marzuki, A. G., Muthmainnah, M., Nursyam, N., Santiana, S., & Erizar, E. (2022). Utilizing Smartphone-Based Pinterest Applications in Developing EFL Students Speaking Skills in Indonesia. *JEES (Journal of English Educators Society)*, 7(2), 205–209. <https://doi.org/10.21070/jees.v7i2.1672>
- Nugroho, A., Putro, N. H. P. S., & Syamsi, K. (2023). The Potentials of ChatGPT for Language Learning: Unpacking its Benefits and Limitations. *Register Journal*, 16(2), 224–247. <https://doi.org/10.18326/register.v16i2.224-247>
- Pikhart, M. (2020). Intelligent information processing for language education: The use of artificial intelligence in language learning apps. *Procedia Computer Science*, 176, 1412–1419. <https://doi.org/10.1016/j.procs.2020.09.151>
- Ratnaningsih, D., Buchari, S. T. H., Fitriani, R., & Rizaldi, M. D. (2025). Artificial Intelligence (AI) in Maritime English Learning Media. 4(2), 128–134. <https://proceeding.unmuhjember.ac.id/index.php/iss/article/view/774>
- Rojabi & Femilia. (2023). EFL learners' perception and attitude in synchronous meetings through Zoom videoconferencing. *JEES (Journal of English Educators Society)*, 8(1), 11–23. <https://doi.org/10.21070/jees.v8i1.1722>
- Thi Thuy Ha, B., & Huu Duc, P. (2024). Difficulties In Teaching and Learning Speaking Skills for Lower Secondary Students at ABC School. *IOSR Journal of Humanities and Social Science (IOSR-JHSS)*, 29(8), 41. <https://doi.org/10.9790/0837-2908044172>
- Van Beuningen, C. G. (2011). The effectiveness of comprehensive corrective feedback in second language writing. *Accelerating the World's Research*, January 2011.
- Yin, J., Khan, R. U., Afzaal, M., Almalki, H. M., Saleem Khasawneh, M. A., & Sulaie, S. Al. (2025). Quantitative risk assessment of speech acts and lexical factors in maritime communication failures and accidents. *Safety Science*, 191(August), 106968. <https://doi.org/10.1016/j.ssci.2025.106968>
- Zhang, C., Meng, Y., & Ma, X. (2024). Artificial intelligence in EFL speaking: Impact on enjoyment, anxiety, and willingness to communicate. *System*, 121(February), 103259. <https://doi.org/10.1016/j.system.2024.103259>

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright ©2025 Dyah Ratnaningsih. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Exploring peer review as a pedagogical tool to enhance paragraph-writing skills in university students

Aidil Syah Putra*,¹ Abdul Rohim,¹ Noor Azlan Ahmad Zanzali,²

¹Universitas Muhammadiyah Tangerang, Indonesia, ²Universiti of Muhammadiyah Malaysia, Malaysia

This study explores the role of structured peer review as a pedagogical tool to enhance paragraph-level writing skills among university students in Indonesia. Academic writing remains one of the most demanding competencies for learners, particularly in producing cohesive and coherent paragraphs. In contexts where large classes and limited teacher feedback constrain the learning process, peer review offers a collaborative approach that distributes responsibility for feedback while fostering deeper engagement with writing. Employing a qualitative case study design, this research involved 25 fourth-year students in an English composition course. Data were gathered from multiple sources, including students' writing drafts collected across several peer review cycles, classroom observations, surveys, and semi-structured interviews. This triangulated approach allowed for a rich exploration of how students' paragraph writing developed over time and how they perceived the peer review process. The findings indicate that students demonstrated notable progress in crafting clearer topic sentences, elaborating supporting details, and producing more consistent concluding sentences. Beyond textual improvements, students reported increased confidence, greater accountability toward peers, and stronger reflective habits. These gains were facilitated by scaffolding strategies such as rubrics, checklists, and teacher modeling, which guided students to provide feedback that moved beyond surface-level corrections toward higher-order concerns of coherence and unity. The study concludes that structured peer review supports not only the improvement of writing products but also the cultivation of collaborative and self-regulated learning practices. As such, it underscores the potential of peer review to be integrated as a core component of English as a Foreign Language (EFL) writing instruction in higher education.

OPEN ACCESS

ISSN 2503 3492 (online)

*Correspondence:

Aidil Syah Putra

aidilisyahputra@umt.ac.id

Received: 6th August 2025

Accepted: 10th September 2025

Published: 31th October 2025

Citation:

Putra, A.S., Rohim, A., Zanzali, N.A.A. (2025).

Exploring peer review as a pedagogical tool to

enhance paragraph-writing skills in university

students. *JEES (Journal of English Educators*

Society), 10(2).

<https://doi.org/10.21070/jees.v10i2.1965>

Keywords: academic writing, higher education, paragraph writing, peer review, writing skills, qualitative case study

INTRODUCTION

Academic writing is a cornerstone of higher education as it enables students to engage deeply with course material, demonstrate their understanding, and develop critical thinking skills (Geithner, A. C. & Pollastro, N. A., 2015). The writing process is fundamental in this endeavor, allowing students to sharpen their ideas, structure their thoughts, and express themselves with clarity and coherence. It requires meticulous planning, revision, and critical self-evaluation to produce well-crafted academic work. Given its complexity, academic writing is often considered one of the most challenging skills for university students to master,

frequently leading to reluctance among some to engage with writing-intensive courses (Dewi et al., 2024).

Paragraph writing is a fundamental component in university-level writing instruction. (Rohim, A., 2019). A well-developed paragraph typically includes a focused topic sentence, supporting details, and a concluding sentence that reinforces the main idea (Siddiqui, A. K., Abbasi, H. R. & Soomro, A., 2023). Mastery of paragraph structure contributes significantly to the development of logical flow, coherence, and the appropriate use of grammar and vocabulary. However, many students still struggle to produce cohesive and focused paragraphs. This struggle often manifests as underdeveloped arguments, disjointed ideas, or a lack of clear progression within their written assignments (Dewi et al., 2024). Such difficulties highlight the necessity for pedagogical approaches that actively engage students in refining their writing processes, moving beyond traditional instruction to more interactive and reflective practices (Trimbur et al., 2001).

In the Indonesian higher-education context, these challenges are compounded by structural factors. Large class sizes and limited opportunities for individualized teacher feedback often prevent students from receiving the level of guidance necessary to refine their writing. This highlights the need for pedagogical approaches that not only strengthen students' paragraph-writing skills but also distribute responsibility for feedback more equitably across the classroom.

Peer review has emerged as a promising pedagogical tool to address these challenges. Through peer review, students exchange feedback, critically analyze each other's work, and improve their understanding of writing conventions (Rahman, 2022; Parr & Timperley, 2010). Studies have shown that peer review enhances paragraph coherence, organization, and clarity by encouraging students to reflect on their writing processes and engage in collaborative learning (Cho & MacArthur, 2009; Coit, 2004; Tsui & Ng, 2000). This interactive approach not only aids in identifying areas for improvement but also fosters a deeper comprehension of rhetorical strategies and grammatical precision (Mallia, 2017).

This collaborative activity reinforces students' grasp of essential paragraph components, such as topic sentences, supporting evidence, and concluding remarks (Cho & MacArthur, 2010).

Prior studies also suggest that peer review fosters metacognitive awareness, independent learning, and student confidence in revising written work (Feng, 2023; Cho & Schunn, 2004; Lundstrom & Baker, 2009). It allows learners to internalize academic standards and develop self-assessment skills. Moreover, feedback from peers often complements or even surpasses instructor comments in improving writing quality. These benefits underscore the value of integrating peer review into writing pedagogy, particularly given the challenges students face in achieving effective written communication (Cho & MacArthur, 2009).

Despite these benefits, few studies have explored the specific effects of peer review on paragraph-level writing, particularly in the context of higher education. Much of the existing literature focuses on broader essay structures or

general writing improvement (Yalch, Vitale, & Ford, 2019; Jegerski & Ponti, 2014). The limited attention to paragraph-focused outcomes leaves a gap in our understanding of how peer feedback supports the development of foundational writing units.

Therefore, this study aims to explore the effectiveness of structured peer review in improving paragraph-level writing skills among university students. By focusing on paragraph elements such as topic sentences, supporting details, and concluding sentences, this study addresses a critical gap in the literature and contributes new insights into how peer review can be implemented to enhance academic writing instruction at the tertiary level.

To guide this investigation, the following research questions were formulated:

1. How does engaging in structured peer review activities influence the paragraph-level writing skills of higher-education students?
2. What specific aspects of paragraph writing (topic sentences, supporting details, and concluding sentences) are most affected by the peer review process?
3. What factors (student perceptions and instructional strategies) contribute to the effectiveness of peer reviews in enhancing paragraph-level writing skills?

METHODS

Research Design

This study employed a qualitative case study design to investigate the effects of peer review on paragraph-level writing skills among higher education students. A case study approach was selected because it allows for an in-depth and contextually grounded exploration of how peer review participation shapes students' ability to produce well-structured and coherent paragraphs within an authentic classroom setting. (Ebneyamini & Moghadam, 2018). Conducted over a semester-long English composition course, the study integrated peer review systematically into writing assignments, enabling the documentation of both observable changes in student writing and students' own reflections on the process.

The peer review intervention was implemented across three review cycles over a 14-week semester. Each cycle included four stages: (1) initial drafting, (2) peer feedback using the provided rubric, (3) self-revision, and (4) teacher confirmation. Classroom observations were conducted five times to examine how students interacted, negotiated meaning, and applied peer feedback. In addition, ten students were purposively selected for semi-structured interviews to represent varying levels of writing performance.

Participant

The participants consisted of 25 undergraduate students enrolled in a fourth-year English composition course at a private university in Indonesia. All participants had previously completed an introductory course in paragraph writing, ensuring a shared foundation in paragraph structure, topic sentence development, supporting detail organization, and coherence techniques. As is common in Indonesian

higher education, the class was relatively large, and opportunities for individualized teacher feedback were limited. This context made peer review a particularly relevant pedagogical strategy, allowing feedback responsibilities to be distributed among students while still strengthening paragraph-writing competencies. Informed consent was obtained from all participants, and ethical approval was secured from the university's research ethics committee.

Data Collection

Data were gathered from multiple sources to provide a comprehensive and triangulated understanding of the phenomenon. The three primary techniques included:

1. Analysis of student writing samples

Writing samples were collected at two points: before the structured peer review intervention and after several cycles of peer review. These samples provided evidence of development in paragraph writing, including changes in structure, coherence, unity, and elaboration of supporting ideas. The comparison of pre- and post-intervention samples offered valuable insights into the extent of improvement in students' paragraph writing.

2. Observation of peer review sessions

Peer review sessions were observed throughout the semester to document how students engaged in the process of giving and receiving feedback. The observations focused on the nature of peer comments (e.g., surface-level corrections vs. content-oriented suggestions), the interactional patterns between reviewers and writers, and the strategies students used to respond to peer feedback in their subsequent drafts. Field notes and observation protocols were used to systematically capture these dynamics.

3. Surveys and interviews with participants

At the end of the course, surveys and semi-structured interviews were conducted to explore students' reflections and experiences with peer review. The surveys captured general perceptions of peer review's effectiveness, while the interviews provided richer, narrative accounts of the perceived benefits, challenges, and learning processes. These instruments added a personal and reflective dimension to the study, highlighting students' voices in relation to their experiences.

The instruments used consisted of a peer review rubric and checklist adapted from remarks ([Cho & MacArthur, 2010](#)), covering clarity of topic sentences, elaboration of supporting details, and paragraph unity. Example items included: "Does the paragraph have one clear focus?" and "Are supporting sentences logically connected to the topic?" Sample interview questions were: "What kind of peer feedback helped you the most?" and "How did peer comments influence your revision decisions?" These instruments ensured consistency and transparency in the data collection process.

The use of multiple data sources ensured methodological triangulation strengthening the credibility of the study by allowing the researchers to examine peer review from different perspectives: textual, observational, and experiential ([Oliver-Hoyo & Allen, 2006](#)).

Data Analysis

Data analysis followed the logic of a qualitative case study, emphasizing contextualized and in-depth understanding. Student writing samples were assessed using a rubric adapted from established criteria for paragraph-level writing, providing descriptive evidence of improvement in coherence, cohesion, unity, grammar, and overall effectiveness. Observational data, survey responses, and interview transcripts were analyzed using manual coding procedures. Open coding was first employed to identify recurring ideas and key concepts, followed by axial coding to group them into broader categories, and selective coding to develop overarching themes. This iterative process enabled close engagement with the data and supported the triangulation of textual, observational, and experiential evidence. The integration of rubric-based analysis with thematic coding provided a holistic narrative of how peer review functioned as a pedagogical tool to enhance paragraph-level writing skills.

RESULTS AND DISCUSSION

The findings in this section are derived from authentic classroom data, including students' paragraph drafts, observation notes, and written reflections collected during the semester. All examples and comments cited here are based on genuine student work rather than simulated or generated data.

Before discussing each research question in detail, it is important to highlight the overall improvement observed in students' paragraph writing. At the beginning of the course, most students wrote short and general paragraphs with vague topic sentences and underdeveloped supporting details. After several peer review cycles, their writing became more organized, focused, and elaborated.

To illustrate this improvement, the following examples show writings taken from students' drafts before and after peer feedback.

Before peer review: "*My hometown is big and noisy.*"

After peer review: "*My hometown is a busy city with crowded streets and many vehicles, but I like living here because people are friendly.*"

Before peer review: "*Pollution is bad.*"

After peer review: "*Pollution from cars and factories makes the air dirty and causes health problems for many people in the city.*"

These examples represent the typical progress made by many participants, who learned to add more specific information and connect their ideas more coherently.

Overall, most students demonstrated improvement in at least one key area of paragraph development—clarity of topic sentences, elaboration of supporting details, or completeness of concluding sentences.

RQ1: Influence of Structured Peer Review on Paragraph-Level Writing Skills

Growth in Overall Writing Competence

The findings revealed that structured peer review had a substantial impact on students' paragraph-level writing competence. By comparing pre- and post-intervention writing samples, it became evident that students made marked

progress in organizing ideas, achieving coherence, and producing unified paragraphs. At the beginning of the semester, many drafts showed weak internal connections; topic sentences were often vague, supporting details were underdeveloped, and paragraphs sometimes contained multiple unrelated ideas. By the end of the course, however, most students demonstrated clearer topic sentences, more elaborated supporting details, and stronger closure through concluding sentences.

Several students attributed these improvements directly to peer review. One participant reflected:

“Before the peer review, I didn’t know my sentences were not connected. My friend told me to make it more related, and I understood my mistake.” (Student 3)

Another emphasized the motivational effect of knowing peers would read their work:

“I tried to write better because my classmates would read my paragraph. I didn’t want to make it messy.” (Student 14)

These accounts highlight how peer review fostered not only skill development but also a sense of accountability and audience awareness, both of which encouraged higher-quality writing.

Development of Reflective Habits

Beyond structural improvements, students also developed reflective writing habits. By engaging in the dual roles of writer and reviewer, they learned to approach their texts more critically. Several described how reviewing peers’ paragraphs led to greater self-awareness. For instance, Student 8 stated:

“When I corrected my friend’s paragraph, I found we made the same mistake. Then I changed my paragraph too.”

Such reflections suggest that peer review operates as a cognitive mirror: by evaluating others, students internalize strategies that they can transfer to their own writing. This process supports the idea that peer review is not just about receiving feedback but also about learning through giving feedback.

Progress Across the Semester

Observational data further showed that the effectiveness of peer review increased over time. In early sessions, students tended to focus on surface-level features such as grammar and word choice. Comments like “check your spelling” or “this sentence is too long” dominated initial feedback sheets. However, as students became more familiar with the structured checklists and gained confidence, their feedback shifted toward higher-order concerns. By mid-semester, they were consistently commenting on idea development, paragraph unity, and logical flow. For example, one mid-semester peer comment read:

“You have two ideas in this paragraph. Try separating them so that the topic sentence is clear.”

This progression suggests that structured peer review is a developmental process: with repeated practice and scaffolding, students learn to engage with deeper dimensions of writing quality.

RQ2: Specific Aspects of Paragraph Writing Most Affected by Peer Review

Improvements in Topic Sentences

Among the three paragraph components examined, topic sentences were the most visibly affected. Early drafts often contained first sentences that were either too general or disconnected from the paragraph’s content. Peer reviewers frequently highlighted this weakness, prompting students to revise.

For example, Student 7 admitted:

“My friend told me that my first sentence was not clear. After that, I tried to write topic sentences that really show what I want to say.”

By the end of the semester, most paragraphs began with clear, purposeful topic sentences that effectively framed the discussion.

Strengthening Supporting Details

Another area of significant growth was the development of supporting details. Initially, many students provided vague or underdeveloped evidence. Peer reviewers often flagged this problem, using comments such as “your example is not specific enough” or “explain this more.” In response, students began incorporating more concrete elaboration into their writing.

For instance, Student 12 revised a paragraph originally written with a single general claim (“Pollution is a big problem in my city”) into one that included specific supporting evidence, such as descriptions of traffic congestion and air quality statistics. The student explained:

“My peer told me my supporting sentences were too general. I revised them by adding examples about the smoke from vehicles and factories.”

This shift reflects a deeper understanding of paragraph development, where claims are supported by relevant and specific details rather than abstract generalizations.

Growth in Concluding Sentences

Although less frequently discussed by students at first, concluding sentences emerged as another key area of improvement. Early drafts often ended abruptly or without synthesis. Through repeated peer review, students became more aware of the need for closure. Peers often wrote comments such as “add a sentence to finish your idea” or “connect back to your topic sentence.”

One student remarked:

“Before, I just stopped writing. My peer said I should write one more sentence to finish the paragraph. So I wrote a conclusion.” (Student 18)

By the final submissions, many paragraphs included concluding sentences that reinforced the topic sentence and synthesized the supporting ideas. Although still an area of challenge for some, the collective progress demonstrated that peer review encouraged greater attention to this overlooked but crucial component of paragraph unity.

Integration of All Three Components

Overall, the peer review process helped students view paragraph writing not as a string of disconnected sentences but as a unified structure with a clear beginning, middle, and

end. This holistic perspective was evident in both student reflections and improved drafts. As Student 20 explained:

“Before, I only thought about grammar. Now I think about how my paragraph starts, develops, and ends.”

RQ3: Factors Contributing to the Effectiveness of Peer Review

Student Perceptions and Attitudes

A critical factor influencing the success of peer review was students’ perceptions of its value. Most participants expressed positive attitudes, emphasizing that peer feedback felt more immediate and less intimidating than teacher comments. Several noted that learning from peers’ mistakes was just as valuable as correcting their own.

One participant reflected:

“When my friend told me my paragraph was confusing, I didn’t feel shy. I just wanted to fix it.” (Student 11)

Positive perceptions fostered higher engagement and a willingness to revise. Conversely, a few students initially expressed skepticism, doubting their peers’ ability to provide useful input. However, as the semester progressed and they observed tangible improvements in their writing, even these students acknowledged the value of the process.

Role of Instructional Strategies

Equally important was the role of structured instructional support. Peer review in this study was not left to chance; it was carefully scaffolded through checklists, rubrics, and teacher modeling. Students consistently emphasized that these tools gave them confidence to provide constructive feedback.

As one participant explained:

“The checklist helped me to know what to check. Before, I didn’t know what to say.” (Student 5)

The teacher also gave examples of how to give feedback politely and clearly. This made students more confident when discussing their writing. By mid-semester, most students were confidently applying the rubric categories—topic sentences, supporting details, and concluding sentences—when reviewing their peers’ work.

Challenges and Limitations

Despite the overall positive outcomes, several challenges were observed. Some students felt hesitant to critique peers strongly for fear of causing offense. Others noted time constraints, particularly during in-class peer review sessions. Nevertheless, these obstacles were mitigated by the structured nature of the activity. Explicit instructions, respectful feedback norms, and the supportive classroom climate encouraged students to overcome initial discomfort.

Importantly, the findings suggest that peer review’s effectiveness depends on the interplay between student perceptions and instructional strategies. Positive student engagement, combined with clear scaffolding, created an environment in which feedback was both meaningful and actionable.

Influence of Peer Review on Paragraph-Level Writing Skills (RQ1)

The first research question sought to understand how structured peer review activities influenced students’

paragraph-level writing skills. The findings revealed notable improvements in students’ ability to construct paragraphs with greater clarity, organization, and logical flow. These gains were most evident in the enhanced quality of topic sentences, the stronger use of supporting details, and the greater consistency in concluding sentences. Beyond the textual features, peer review also fostered metacognitive awareness, as students became more attentive to the conventions of academic writing and more reflective about their own writing processes.

These findings align with earlier studies that underscore the positive role of peer feedback in promoting writing development (Cao et al., 2022; Nguyen, 2016). Peer review provided students with opportunities to critically evaluate the work of their classmates, which in turn encouraged them to internalize standards of effective writing and transfer those insights into their own revisions. This reciprocal process resonates with Vygotskian perspectives on learning, particularly the notion of the zone of proximal development (ZPD), where learners achieve more with the guidance of peers than they might independently. In this sense, peer review did not function merely as a corrective mechanism but as a dialogic and collaborative learning tool that scaffolded students toward greater proficiency.

Importantly, the study highlights that peer review is particularly effective when structured and guided by clear criteria. Students reported that using rubrics and checklists helped them focus their feedback and reduced the anxiety of critiquing peers’ work. This corroborates findings from (Cho & Cho, 2010) who noted that structured peer review fosters higher-quality feedback and more substantial revisions. In the current study, structured peer review not only shaped students’ technical skills in writing but also fostered greater confidence in their ability to evaluate and produce academic text.

Specific Aspects of Paragraph Writing Affected by Peer Review (RQ2)

The second research question addressed which specific aspects of paragraph writing were most affected by the peer review process. The findings indicated that while all three aspects (topic sentences, supporting details, and concluding sentences) showed improvement, the most substantial gains occurred in the use of supporting details. Students became more adept at providing examples, evidence, and explanations that substantiated their topic sentences. This suggests that peer review was particularly effective in moving students beyond surface-level writing toward deeper elaboration and argumentation.

The improvement in supporting details can be attributed to the dialogic nature of peer review, where students were prompted to ask for clarification, request elaboration, or challenge weak arguments in their peers’ writing. This interaction mirrored authentic academic discourse and encouraged students to consider the expectations of an audience. Previous research has emphasized the importance of audience awareness in academic writing and the present findings extend this work by illustrating how peer review cultivates that awareness at the paragraph level (Baker, 2016; Caux & Pretorius, 2024). This collaborative knowledge-

building process, where students engaged in providing feedback, significantly contributed to their ability to develop and articulate comprehensive supporting arguments within their paragraphs (Lim & Tay, 2024). This active engagement in providing feedback also cultivated self-assessment skills, as students began to apply similar evaluative criteria to their own work (Bergamin et al., 2019).

Topic sentences also showed significant improvement, with students demonstrating a clearer ability to state the main idea of each paragraph concisely and appropriately. Peer reviewers frequently identified vague or overly broad topic sentences, prompting writers to revise them for precision and focus. This reflects the findings of (Baker, 2016; Lundstrom & Baker, 2008), who argue that peer feedback strengthens writers' awareness of rhetorical structure and coherence. In the present study, peer feedback guided students toward recognizing the central role of topic sentences in paragraph unity.

Concluding sentences exhibited moderate improvement compared to topic sentences and supporting details. Many students still struggled to provide effective summaries or reflective statements at the end of paragraphs. While peer review raised awareness of this weakness, fewer comments directly addressed concluding sentences, suggesting that this aspect of paragraph writing may require more explicit instructional intervention. This finding underscores the need for teachers to integrate targeted instruction on paragraph closure into peer review activities, ensuring that students fully grasp the rhetorical purpose of concluding sentences.

Factors Contributing to the Effectiveness of Peer Review (RQ3)

The third research question focused on the factors that contributed to the effectiveness of peer review, with particular emphasis on student perceptions and instructional strategies. Several key factors emerged.

Student Perceptions

Students overwhelmingly viewed peer review as beneficial for their writing development. They appreciated the opportunity to receive immediate and varied feedback, which they often described as more relatable and less intimidating than teacher feedback. Many students also reported that reviewing others' work helped them recognize flaws in their own writing, reinforcing the dual benefits of giving and receiving feedback. This supports findings by (Nicol et al., 2013), who argue that peer assessment enhances students' evaluative judgment and self-regulated learning.

However, the study also uncovered challenges in student perceptions. Some students initially doubted their ability to provide useful feedback, expressing concern about their own limited proficiency. Over time, the provision of structured rubrics and guided practice alleviated these anxieties, enhancing students' confidence and willingness to engage. This suggests that perceptions of self-efficacy play a critical role in shaping the success of peer review.

Instructional Strategies

Instructional design was another critical factor. The effectiveness of peer review depended heavily on the teacher's ability to provide clear guidelines, scaffolding, and follow-up. Rubrics and checklists proved indispensable in

helping students focus on the key aspects of paragraph writing. Moreover, teacher modeling of effective feedback demonstrated the expected tone, specificity, and constructiveness of peer comments. Without such instructional scaffolds, peer review risked devolving into superficial praise or unhelpful critique.

Another instructional factor was the integration of peer review into a cyclical process of drafting, feedback, and revision. Students recognized that peer review was most valuable when it directly informed subsequent revisions rather than functioning as an isolated activity. This aligns with (Manchón & Matsuda, 2016), who emphasizes the iterative nature of writing development and the importance of multiple opportunities for feedback and revision. The current findings affirm that peer review must be embedded within a broader pedagogical framework that prioritizes writing as a process rather than a product.

Social and Cultural Dynamics

Finally, social and cultural dynamics influenced the peer review process. In some cases, students were reluctant to provide critical feedback to peers for fear of causing offense, particularly in collectivist cultural contexts where maintaining group harmony is valued (Carson & Nelson, 1994). Teachers addressed this by emphasizing the constructive purpose of feedback and by framing peer review as a collaborative effort toward mutual improvement. Over time, students became more comfortable offering critical yet respectful suggestions, highlighting the importance of cultural sensitivity in peer review implementation.

Theoretical Implications

The findings of this study contribute to theoretical discussions on peer review and writing pedagogy in several ways. First, they reinforce sociocultural perspectives on learning, illustrating how writing development is mediated through interaction with peers and guided by shared tools such as rubrics. The observed improvements in paragraph structure confirm that learning is not merely individual but co-constructed through social processes. This resonates with Vygotsky's claim that higher mental functions develop first on the social plane before becoming internalized.

Second, the study advances our understanding of peer review as both a cognitive and affective process. Cognitively, students developed greater awareness of paragraph structure and rhetorical clarity. Affectively, they gained confidence, motivation, and a sense of belonging within the writing community. This dual dimension highlights the holistic impact of peer review, suggesting that its benefits extend beyond textual outcomes to encompass broader developmental goals in higher education.

Pedagogical Implications

From a pedagogical perspective, the findings suggest several implications for writing instruction in higher education. First, structured peer review should be integrated as a regular feature of writing courses, supported by clear rubrics and guided practice. Teachers should emphasize not only how to give feedback but also how to use it effectively in revision. Second, special attention should be devoted to aspects of writing that are less frequently addressed in peer feedback, such as concluding sentences, to ensure balanced

development of all components of paragraph structure. Third, teachers should foster a classroom culture that values constructive critique, mutual respect, and collaborative learning, thereby reducing anxiety and resistance to peer review.

The study also highlights the potential of peer review to complement teacher feedback. While teacher feedback remains essential for addressing complex linguistic and rhetorical issues, peer feedback provides immediacy, variety, and opportunities for critical engagement that teacher feedback alone cannot fully replicate. Together, these forms of feedback create a more comprehensive and supportive environment for writing development.

Limitations and Directions for Future Research

Although this study provides valuable insights, certain limitations must be acknowledged. The sample size was relatively small and context-specific, which may limit the generalizability of findings to broader populations. Future research should replicate the study with larger and more diverse cohorts across different institutional and cultural contexts. Additionally, the study relied primarily on qualitative data from observations and interviews, complemented by textual analysis of students' paragraphs. Future work could integrate more longitudinal designs or experimental approaches to examine the sustained effects of peer review on writing development.

Moreover, while this study focused on paragraph-level writing, future research could extend to more complex genres such as essays, reports, or research papers. Investigating whether the benefits of peer review observed at the paragraph level transfer to larger and more sophisticated writing tasks would further enrich our understanding of its pedagogical value.

CONCLUSION

This study investigated how structured peer review activities influence paragraph-level writing skills among university students. The findings confirm that peer review, when designed with clear scaffolding and instructional support, can significantly improve students' abilities to construct coherent and unified paragraphs. It enhances three key elements of paragraph writing—topic sentences, supporting details, and concluding sentences—by fostering critical reflection, reader awareness, and metacognitive engagement.

Three main conclusions can be drawn. First, structured peer review contributes meaningfully to the development of paragraph-level writing by encouraging students to identify and revise structural weaknesses in their work. Second, the most affected components of paragraph writing are those that determine clarity and coherence: the clarity of topic focus, the adequacy of elaboration, and the completeness of closure. Third, the effectiveness of peer review is closely tied to two factors: students' positive perceptions of the activity and the presence of explicit instructional guidance, such as rubrics and checklists. Beyond its impact on writing outcomes, peer review also cultivates student agency, collaboration, and self-evaluation—qualities that are essential in learner-centered pedagogy and reflective academic practice.

Based on these conclusions, several practical and pedagogical recommendations are proposed. Instructors are encouraged to integrate structured peer review into writing curricula not as an add-on, but as a core instructional strategy. To ensure its success, peer review activities should be supported with clear guidelines, feedback forms, and modeling of constructive commentary. Teacher preparation plays a vital role: educators must provide orientation on how to give focused, respectful, and meaningful feedback. Institutions should also consider including peer review frameworks in their academic writing modules, particularly at the paragraph level, which lays the foundation for more advanced writing.

Future research may explore how peer review affects long-term writing development, the role of peer dynamics in shaping feedback quality, or its impact in asynchronous online settings. Moreover, comparative studies across different cultural or disciplinary contexts could enrich our understanding of peer review's pedagogical potential.

ACKNOWLEDGEMENTS

The authors would like to express their sincere gratitude to the students who participated in this study and to the academic staff who supported the implementation of the peer review activities. Special thanks are extended to the advisors and colleagues who provided valuable feedback during the research process. The authors also acknowledge the contributions of those who assisted in proofreading and preparation of the manuscript. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

- Baker, K. M. (2016). Peer review as a strategy for improving students' writing process [Review of Peer review as a strategy for improving students' writing process]. *Active Learning in Higher Education*, 17(3), 179. SAGE Publishing. <https://doi.org/10.1177/1469787416654794>
- Bergamin, P., Bosch, C., Toit, A. D., Goede, R., Golightly, A., Johnson, D. W., Johnson, R. T., Kruger, C., Laubscher, D., Lubbe, A., Olivier, J., Westhuizen, C. van der, & Zyl, S. van. (2019). Self-Directed Learning for the 21st Century: Implications for Higher Education. In *NWU self-directed learning series*. <https://doi.org/10.4102/aosis.2019.bk134>
- Cao, S., Zhou, S., Luo, Y., Wang, T., Zhou, T., & Xu, Y. (2022). A review of the ESL/EFL learners' gains from online peer feedback on English writing [Review of A review of the ESL/EFL learners' gains from online peer feedback on English writing]. *Frontiers in Psychology*, 13. Frontiers Media. <https://doi.org/10.3389/fpsyg.2022.1035803>
- Carson, J. G., & Nelson, G. L. (1994). Writing groups: Cross-cultural issues. *Journal of Second Language Writing*, 3(1), 17. [https://doi.org/10.1016/1060-3743\(94\)90003-5](https://doi.org/10.1016/1060-3743(94)90003-5)

- Caux, B. C. de, & Pretorius, L. (2024). Learning together through collaborative writing: The power of peer feedback and discussion in doctoral writing groups. *Studies In Educational Evaluation*, 83, 101379. <https://doi.org/10.1016/j.stueduc.2024.101379>
- Cho, K., & MacArthur, C. A. (2009). Student revision with peer and expert reviewing. *Learning and Instruction*, 20(4), 328. <https://doi.org/10.1016/j.learninstruc.2009.08.006>
- Cho, Y. H., & Cho, K. (2010). Peer reviewers learn from giving comments. *Instructional Science*, 39(5), 629. <https://doi.org/10.1007/s11251-010-9146-1>
- Cho, K., & Schunn, C. D. (2004). You Write Better When You Get Feedback from Multiple Peers Than an Expert. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, 26(26). <https://escholarship.org/uc/item/7sx4c3rx>
- Coit, C. (2004). Peer review in an online college writing course. *Proceedings of the IEEE International Conference on Advanced Learning Technologies (ICALT)*. <https://doi.org/10.1109/icalt.2004.1357712>
- Dewi, A. C., Mohamed, S., Rahman, K., & Cardoso, L. M. (2024). The Impact of Performance-Based Assessment (PBA) on Academic Writing Instruction.
- Ebneyamini, S., & Moghadam, M. R. S. (2018). Toward Developing a Framework for Conducting Case Study Research. *International Journal of Qualitative Methods*, 17(1). <https://doi.org/10.1177/1609406918817954>
- Feng, W., Zhu, W., Fu, T. J., Jampani, V., Akula, A., He, X., & Wang, W. Y. (2023). Layoutgpt: Compositional visual planning and generation with large language models. *Advances in Neural Information Processing Systems*, 36, 18225-18250. https://proceedings.neurips.cc/paper_files/paper/2023/hash/3a7f9e485845dac27423375c934cb4db-Abstract.html
- Geithner, C. A., & Pollastro, N. A. (2015). Constructing engaged learning in scientific writing. *Journal of Applied Research in Higher Education*, 7(2), 292–307. <https://doi.org/10.1108/jarhe-04-2014-0053>
- Jegerski, J., & Ponti, E. (2014). Peer review among students of Spanish as a heritage language: The effectiveness of a metalinguistic literacy task. *Linguistics and Education*, 26, 70-82. <https://doi.org/10.1016/j.linged.2014.03.002>
- Lim, F. V., & Tay, M. Y. (2024). Multiliteracies & Multimodality: A Classroom Inquiry Study in Singapore.
- Lundstrom, K., & Baker, W. (2008). To give is better than to receive: The benefits of peer review to the reviewer's own writing. *Journal of Second Language Writing*, 18(1), 30. <https://doi.org/10.1016/j.jslw.2008.06.002>
- Mallia, J. (2017). Strategies for Developing English Academic Writing Skills. *Arab World English Journal*, 8(2), 3. <https://doi.org/10.24093/awej/vol8no2.1>
- Manchón, R. M., & Matsuda, P. K. (2016). Handbook of Second and Foreign Language Writing. In De Gruyter eBooks. De Gruyter. <https://doi.org/10.1515/9781614511335>
- Nguyen, H. T. (2016). Peer Feedback Practice in EFL Tertiary Writing Classes. *English Language Teaching*, 9(6), 76. <https://doi.org/10.5539/elt.v9n6p76>
- Nicol, D., Thomson, A., & Breslin, C. (2013). Rethinking feedback practices in higher education: a peer review perspective [Review of Rethinking feedback practices in higher education: a peer review perspective]. *Assessment & Evaluation in Higher Education*, 39(1), 102. Taylor & Francis. <https://doi.org/10.1080/02602938.2013.795518>
- Oliver-Hoyo, M. T., & Allen, D. (2006). The Use of Triangulation Methods in Qualitative Educational Research. *The Journal of College Science Teaching*, 35(4), 42. <https://faculty.yu.edu.jo/Audeh/My%20Gallery/paper%20and%20documents/qualitative%20paper4.pdf>
- Parr, J. M., & Timperley, H. (2010). Feedback to writing, assessment for teaching and learning and student progress. *Assessing Writing*, 15(2), 68–85. <https://doi.org/10.1016/j.asw.2010.05.004>
- Rahman, M. M. (2022). The essentials of paragraph writing to develop writing skill. *Global Journal of Human-Social Science: G Linguistics & Education*, 22(5), 71–75. <https://doi.org/10.34257/gjhssgvol22is5pg71>
- Rohim, A. (2019). An analysis of students writing skill in paragraph writing. *Globish (An English-Indonesian Journal for English, Education and Culture)*, 8(1), 40–48. <http://dx.doi.org/10.31000/globish.v7i2.1126>
- Siddiqui, A. K., Abbasi, H. R., & Soomro, A. (2023). Paragraph organization errors in the writing of Pakistani college-going students: An error analysis study. *Asian EFL Studies and Scholarly Review*, 3(2), 131–139. <https://doi.org/10.48112/aessr.v3i2.479>
- Trimbur, J., Cope, B., & Kalantzis, M. (2001). Multiliteracies: Literacy Learning and the Design of Social Futures. *College Composition and Communication*, 52(4), 659. <https://doi.org/10.2307/358703>
- Tsui, A. B. M., & Ng, M. (2000). Do secondary L2 writers benefit from peer comments? *Journal of Second Language Writing*, 9(2), 147–170. [https://doi.org/10.1016/s1060-3743\(00\)00022-9](https://doi.org/10.1016/s1060-3743(00)00022-9)
- Yalch, M. M., Vitale, M. E., & Ford, K. J. (2019). Benefits of peer review on students' writing. *Active Learning in Higher Education*, 18(3), 317–325. <https://doi.org/10.1177/1475725719835070>

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright ©2025 Aidil Syah Putra, Abdul Rohim, Noor Azlan Ahmad Zanzali. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Creative teaching in English language classrooms: A reflective and contextual interpretation

Muhamad Ahsanu*, **Erna Wardani**, **Ashari Ashari**, **Nisa Roiyasa**
Universitas Jenderal Soedirman, Indonesia

Creativity has become a defining competency in twenty-first-century higher education, particularly in English language teaching (ELT), yet how university educators conceptualize and enact it remains under-explored. This study investigates the lived experiences of two Indonesian lecturers—one from a public and one from a private university—through a narrative inquiry approach supported by classroom observations and thematic analysis. The findings reveal that creative teaching in ELT is a reflective, adaptive, and contextually grounded process characterized by continual negotiation between pedagogical innovation and institutional constraints. Two interrelated models emerged: the Five-Facet Framework (FFF) (teaching capital, material, media, method, and evaluation) and Reflective Adaptation as the sustaining mechanism of creativity. Effective creative teaching was found to rely not on isolated techniques but on the orchestration of these five components, harmonized like instruments in an orchestra. Teachers' passion, autonomy, and responsiveness to learners' needs serve as catalysts that transform routine instruction into dynamic, student-centered learning experiences. Reflection and contextual empowerment further enable resilience amid limited resources and rigid curricula. The study contributes to sociocultural and constructivist perspectives by framing creativity as both a personal and collective endeavour—rooted in interaction, reflection, and contextual engagement. Theoretically, it reaffirms that creative teaching forms the foundation of effective pedagogy; practically, it advocates institutional support systems that nurture reflective professional cultures and empower teachers to sustain innovation in ELT.

Keywords: creative teaching, creativity in ELT, higher education, reflective pedagogy, narrative inquiry, contextual innovation

OPEN ACCESS

ISSN 2503 3492 (online)

*Correspondence:

Muhamad Ahsanu

muhamad.ahsanu@unsoed.ac.id

Received: 13th August 2025

Accepted: 16th September 2025

Published: 31th October 2025

Citation:

Ahsanu, M., Wardani, E., Ashari, A., Roiyasa, N. (2025). Creative teaching in English language classrooms: A reflective and contextual interpretation. *JEEES (Journal of English Educators Society)*, 10(2). <https://doi.org/10.21070/jees.v10i2.1967>

INTRODUCTION

In contemporary higher education (HE), the integration of creativity into pedagogical practices has become an increasingly prominent focus, particularly within the context of English language instruction. The infusion of creativity in teaching, often referred to as Creativity in Language Teaching (CiLT), is now regarded as essential rather than supplementary. As higher education institutions deal with escalating demands driven by globalization, digital innovation, and the shift toward more engaging instructional systems, creative pedagogy (CP) has gained increasing recognition as a transformative educational model. Its rising relevance is evident in university reform agendas and contemporary pedagogical scholarship, both of which emphasize innovation and pedagogical flexibility over traditional, transmissive models of instruction (Heidari-Shahreza, 2024). In this context, the present study approaches creativity

not solely as an individual cognitive ability but as a deliberate, strategic, and context-responsive teaching practice situated within Indonesian tertiary English language education. This focus reflects the urgent imperative to redesign language instruction in ways that address the complex intellectual, social, and affective demands of 21st-century learners.

English language educators frequently negotiate the dual pressures of maintaining established instructional standards while simultaneously adopting more exploratory and student-centered approaches. Although such negotiations can be challenging, they also provide conducive milieu for the creation of learning environments that are more participatory, culturally relevant, and personally meaningful. Creative pedagogical approaches enable instructors to respond flexibly to different learner profiles and emerging classroom dynamics. However, the continuance of conventional teaching methods in many higher education settings continues to draw critique from scholars who argue that rigid, test-oriented practices disproportionately benefit analytically dominant learners while constraining those with creative strengths (Sternberg, 2006; Ismayilova & Laksov, 2022). Against this backdrop, the current research places attention on educators who actively destabilize prevailing norms by incorporating creative strategies into their pedagogical planning. Specifically, it examines how English language lecturers conceptualize, operationalize, and sustain creative pedagogy in their everyday instructional decision-making processes. By centering on practitioners of understanding how creativity is operationalized within authentic classroom contexts.

The central premise of this investigation is built around a fundamental proposition: Creative teaching in English language classrooms hinges on a reflective and contextual interpretation. Exploring this fundamental construct in language teaching requires a multidimensional approach that examines both the characteristics of creative instruction and its tangible impact on students and teachers. Creativity in teaching is not merely about entertainment or aesthetic sense; it is closely linked to instructional effectiveness. By identifying the specific elements that constitute creative pedagogy in higher education (HE), this study seeks to clarify how such practices contribute to deeper student engagement and improved academic outcomes.

Of particular interest is how faculty members define creativity in pedagogical terms and how these definitions translate into classroom practice. Through this lens, creativity is examined not only as an abstract ideal but also as a functional teaching tool. This line of inquiry lays the foundation for uncovering the essential features of effective creative teaching. Several foundational studies have laid the groundwork for the evolving discourse on creativity in language education. Kettler, Lamb, Willerson, and Mullet (2018) observed that teachers who value creativity are more likely to foster it among their students. Their research suggests that such educators promote student autonomy, critical thinking, and a deeper sense of ownership in the learning process.

Similarly, Ismayilova and Laksov (2022) argue that creativity allows educators to navigate complex classroom dynamics, enhance student participation and introduce

innovative content delivery methods. Collectively, these studies underscore the dual influence of personal attributes and institutional culture in either enabling or constraining creative teaching. Together, they provide a strong foundation upon which the present study is situated. In another pivotal contribution, Yu, Wang, and Yuizono (2023) examined the role of creativity in enhancing student's questioning skills, a key component of critical thinking and linguistic competence. Their findings demonstrated that student-centered creative techniques improved learners' ability to formulate questions, which, in turn, led to greater vocabulary and grammar acquisition. This research reinforces the notion that creativity in teaching is not a luxury but a necessity for language development, particularly in English as a Foreign Language (EFL) contexts.

Nonetheless, the integration of creative practices is not without its challenges. Frawley (2020) highlighted how institutional and structural limitations often hinder teachers from realizing their creative potential, leading to professional fatigue or disillusionment. This underscores the importance of addressing systemic barriers to foster sustainable creative teaching. Richards and Cotterall (2016) outlined two primary benefits of creative instruction: it fosters students' imaginative thinking and builds self-confidence, while simultaneously energizing teachers by enhancing job satisfaction. Building on this, Khany and Boghayeri (2014) argued that HE instructors must not only understand creative teaching techniques but also apply them systematically within their classrooms. Creativity, in this light, is a skill that can and should be developed. It is not merely a personal attribute, but a pedagogical imperative linked to broader societal progress. Elisondo (2016) further emphasized that creativity is central to advancements in science, art, and technology, marking it as a cornerstone of human development and societal transformation.

Such systemic issues underscore the need for localized strategies to embed creativity into English language instruction. Creativity's central role in ELT is widely acknowledged for its positive impact on student motivation, participation, and learner-centered instruction. Studies by Richards and Cotterall (2016) and Kettler et al. (2018) highlight how creative strategies can significantly improve classroom dynamics and learning outcomes.

Despite growing recognition of the pedagogical value of creativity, its implementation in higher education (HE) remains constrained by rigid curricula, bureaucratic structures, and institutional resistance (Frawley, 2020; Ismayilova & Laksov, 2022). One of the enduring challenges in this domain lies in reconciling theoretical conceptions of creative pedagogy with their operationalization in classroom contexts. This study seeks to address this gap by offering empirical insights into how creative teaching is conceptualized and enacted within Indonesian universities.

To this end, the study proposes a functional framework for creative pedagogy emphasizing adaptability, multiple teaching approaches, and responsiveness to sociocultural and institutional learning contexts (Sawyer, 2011). Grounded in key theoretical traditions—namely Holec's (1981) notion of learner autonomy, Vygotsky's (1978) sociocultural theory, and Boden's (1990) models of creativity—the framework

situates teachers' beliefs and practices within broader paradigms of educational innovation. As [Heidari-Shahreza \(2024\)](#) contends, faculty belief systems significantly shape pedagogical choices, with institutional support serving as a crucial mediator that determines whether such beliefs materialize into innovative teaching practices. Conversely, obsolete tools and unsupportive learning environments often stifle even the most motivated educators. The case of a creative lecturers, Shanti and Shinta, illustrates how pedagogical creativity evolves dynamically in response to student participation and feedback ([Khany & Boghayeri, 2014](#)), providing a contextual foundation for analysing creativity in authentic teaching settings.

Creativity in teaching constitutes a multifaceted construct that merges imaginative thinking with pedagogical decision-making to enhance learning engagement and efficacy. [Freire \(2005\)](#) and [Wisdom \(2006\)](#) argued that creativity should form a central component of educational praxis; nevertheless, many HE institutions struggle to actualize this ideal. Divergent understandings of what constitutes creativity among educators contribute to this inconsistency. For some, creativity denotes the generation of novel ideas, while for others, it involves problem-solving or the reinterpretation of existing knowledge and materials ([Beghetto, 2007](#); [Dobbins, 2009](#)). To mitigate such ambiguities, teacher education programs must offer explicit frameworks for embedding creativity across lesson planning, assessment, and classroom interaction ([Rinkevich, 2011](#)). Hence, establishing a shared conceptual and operational understanding of creativity is a prerequisite for its coherent integration into pedagogical practice.

From a cognitive and psychological perspective, creativity is often interpreted through the dual framework of "big-C" and "little-c" creativity ([Kaufman & Sternberg, 2019](#)). The former represents eminent, world-changing innovations, whereas the latter pertains to everyday acts of creative problem-solving and instructional adaptation, such as designing engaging learning activities or reconfiguring teaching materials ([Alves-Oliveira et al., 2021](#)). In educational contexts, most manifestations of creativity correspond to "little-c" creativity, which involves recontextualizing existing knowledge to address immediate instructional needs. [Boden \(2004\)](#) further distinguished creative cognition into combinational, exploratory, and transformational processes, each contributing to the reformation of pedagogical knowledge and practice. These perspectives collectively offer a theoretical scaffold for analyzing and cultivating creativity within classroom environments.

Nevertheless, institutional conventions frequently restrict creative expression in education. Mechanisms such as standardized testing, prescriptive syllabi, and rigid evaluative procedures often deter experimentation and innovation ([Dawson et al., 2011](#); [Mullet et al., 2016](#)). Such institutional rigidity underscores the need for empirical inquiry into how teacher beliefs, professional preparation, and contextual factors interact to shape creative pedagogical behavior. Although many educators express a willingness to innovate, their efforts are often hindered by insufficient resources, limited time, and a lack of institutional encouragement.

Accordingly, this study aims to elucidate the conditions that enable or constrain creativity within higher education language instruction.

Within the Indonesian context, English language teaching (ELT) continues to encounter persistent challenges. Traditional teaching paradigms emphasizing rote memorization and textbook dependency still dominate, thereby restricting opportunities for communicative engagement and critical inquiry ([Puspitasari & Purnomo, 2020](#)). Standardized curricula further curtail the pedagogical space for innovation ([Wahyu, 2021](#)), while disparities in access to professional development and technological resources—particularly in rural regions—compound these issues ([Suyadi & Zakaria, 2021](#); [Amin, 2022](#)). Addressing such contextual impediments is essential for devising sustainable and contextually relevant interventions to foster creativity in language teaching.

Despite these challenges, creativity offers significant potential to revitalize ELT in Indonesia. Empirical evidence indicates that creative educators are more effective in engaging students, stimulating motivation, and cultivating reflective thinking ([Adnan, 2022](#); [Hidayat et al., 2023](#)). They tend to employ student-centered pedagogies such as project-based learning, task-based instruction, and digital media integration to enhance interactivity and contextual relevance in language learning ([Suryati & Suwarno, 2020](#)). These practices demonstrate that creativity, when purposefully cultivated, can function as both a practical and transformative force within the classroom.

In conclusion, creativity in English language teaching within higher education is a complex, context-sensitive, and transformative notion. While its pedagogical benefits are widely recognized across both global and Indonesian contexts, its systematic implementation remains obstructed by structural constraints, outdated pedagogical orientations, and insufficient teacher preparation. This study, therefore, endeavours to investigate how Indonesian university English lecturers conceptualize, apply, and are institutionally supported in their creative teaching endeavours. The insights derived are anticipated to inform the development of more responsive teacher training models, evidence-based policy reform, and enhanced pedagogical practice in higher education language education.

METHODS

Narrative Inquiry

Narrative inquiry is a qualitative research method that centers on collecting data through individuals' personal stories. It embodies a continuous process of experiencing, articulating, revisiting, and reinterpreting one's lived stories ([Clandinin & Connelly, 2000](#)). It positions researchers' personal histories and inquiries as integral to producing authentic knowledge ([Clandinin, 2013](#)). Guided by the recognition that numerous meaningful narratives remain unspoken, this approach further serves as a critical space for amplifying marginalized perspectives ([Creswell, 2012](#)). By emphasizing personal narratives, it deepens our understanding of teaching, learning, and educational performance, and can inform the development of more effective teaching strategies ([Meegan,](#)

2023). As Meegan's notes, narratives shape how educators perceive their learners and offer valuable insights for addressing both theoretical and practical educational challenges.

This study applies narrative inquiry to explore an individual's lived experience. As outlined by [Clandinin and Connelly \(2000\)](#), this approach is deeply contextual and connects moments across time and intentional action, acknowledging the continuity of past, present, and future. We conducted narrative interviews with an English lecturer at a university in Purwokerto, Indonesia. Through her storytelling, the study highlights her experiences in teaching language skills, aiming to share insights with fellow educators. Narrative inquiry fosters collaboration between researcher and participant, generating meaningful, shared understanding ([Clandinin & Connelly, 2000](#); [Wei, 2023](#)).

Sampling Strategy

This study employed purposeful sampling, specifically criteria sampling - a of purposive sampling ([Creswell & Poth, 2018](#)). Purposeful sampling is widely recognized in qualitative research for its ability to identify information-rich cases that can provide in-depth insights into a phenomenon ([Patton, 2015](#)). The participant was selected based on three criteria: (a) having taught language skills in an English department at higher education level, (b) possessing more than five years of teaching experience, and (c) demonstrating a passion for teaching, a willingness to engage in professional development, and enthusiasm for applying new learning in the classroom. This approach aligns with the principle that effective qualitative sampling prioritizes participants who are especially knowledgeable or experienced in the topic of interest ([Merriam & Tisdell, 2016](#)). The participants, referred to hereby the pseudonym Shinta and Shanti, taught content courses in EFL classes in a public university.

Data Collection

This study employed narrative interviews and observation as the primary data collection methods, comprising two different sessions with the participants. The initial session involved an in-depth conversation guided by semi-structured questions, allowing the participant to share her experiences in a conversational and open manner. Follow-up interviews were intentionally structured to assist the participants in refining, elaborating, and articulating their personal accounts with greater precision and depth. Designed in an open-ended, exploratory manner, these sessions facilitated expansive storytelling and ensured the generation of rich, contextually embedded descriptions that capture the multifaceted nature of her lived experiences ([Hamadou, 2024](#)). Each interview lasted between 90 and 120 minutes, allowing ample temporal flexibility and enabling a more participant-driven pace for narrative construction. In addition, systematic classroom observations were conducted to complement the interview data and strengthen the overall narrative dataset.

Data Analysis and Interpretation

The analytical process employed a narrative inquiry approach integrated with thematic analysis, highlighting both the experiential dimension of the participants' stories and the patterned meanings emerging across narrative and

observational sources. Data analysis progressed through a cyclical, reflexive engagement with the empirical material, balancing inductively derived insights with theoretically informed interpretations ([Bingham, 2023](#)). Interview recordings were transcribed verbatim and underwent iterative coding to reconstruct narrative sequences around key events and thematic patterns. Member checking procedures were implemented to uphold validity and authenticity. Continuous re-examination of transcripts and field notes supported deeper immersion into the participant's perspective, enabling the researcher to more comprehensively "experience the experience" ([Clandinin & Connelly, 2000](#); [Proudfoot, 2023](#)).

RESULTS AND DISCUSSION

The findings and discussion unfold in two interconnected parts, encompassing insights from classroom observations and in-depth interviews with two English teachers, Shinta and Shanti. Using thematic analysis, both data sources were examined to capture the recurring patterns and underlying meanings that define creativity in language teaching. The analysis moves beyond description to offer interpretive insights into how creativity emerges and operates in classroom contexts. In particular, the observation results illuminate three pivotal research questions:

1. What distinctive features frame and shape creativity in language teaching?
2. How do teachers translate creative principles into tangible classroom practices?
3. What foundational elements serve as the driving forces behind creative teaching?

The classroom observations conducted in two distinct English language teaching contexts—Shinta's and Shanti's classes—revealed a vibrant interplay among creativity, reflective thought, and contextual responsiveness that shaped their pedagogical practices. Observation data demonstrated that creativity functioned as a consistent pedagogical orientation rather than an occasional or spontaneous action. Teachers regularly infused creative choices into instructional planning and classroom decision-making. The interpretive thematic analysis identified seven key dimensions guiding their creative practice in English language teaching: defining and demonstrating creativity, sources of creative inspiration, inventive instructional strategies, techniques for enhancing student engagement and development, contextual supports and constraints, reflective professional growth, and the creation of a positive classroom climate. These dimensions were consolidated into four interconnected themes and one overarching integrative construct: Sustaining Creativity in Language Teaching through Reflective Adaptation and Contextual Empowerment. This core theme captures creativity as an evolving process in which educators refine approaches through reflection, innovation, and strategic responsiveness to institutional and sociocultural demands to foster meaningful language learning.

Creativity as a Dynamic Pedagogical Process

The observation data underscored that creativity was embedded in the teachers' pedagogical philosophy rather than

being an occasional instructional choice. Their lessons integrated multimodal and participatory approaches such as storytelling, games, role-plays, and contextualized discussions, transforming routine classroom sessions into spaces of exploration and engagement. These strategies exemplified [Richards' \(2013\)](#) idea of “pedagogical creativity,” emphasizing the teacher’s capacity to devise novel and contextually appropriate methods to facilitate meaningful learning.

Both educators strategically integrated visual resources and culturally grounded content reflective of students’ immediate sociolinguistic environments, thereby transforming abstract linguistic concepts into contextually meaningful learning experiences. Through the implementation of multimodal activities—such as interactive games, role-play, and dramatized performance—the teachers not only strengthened learners’ grammatical and lexical development but also enhanced intrinsic motivation, positive affect, and communicative fluency. These pedagogical practices align closely with [Cremin’s \(2015\)](#) characterization of creative language educators as “possibility thinkers,” who intentionally craft learning spaces that nurture curiosity, adaptive thinking, and imaginative engagement within classroom interaction.

Furthermore, the teachers’ questioning styles demonstrated deliberate efforts to stimulate divergent thinking. Instead of convergent, single-answer questioning, they frequently posed open-ended inquiries that encouraged multiple perspectives and linguistic experimentation. This approach aligns with [Sawyer’s \(2011\)](#) concept of “disciplined improvisation,” in which structure and spontaneity coexist, enabling originality within guided frameworks. By embedding humour, playful interaction, and culturally situated content into instructional delivery, the teachers effectively transformed English language lessons into dynamic learning spaces where learners were encouraged to explore language and develop expressive confidence. In this context, creativity emerged not as a supplementary instructional feature but as a foundational pedagogical orientation that shaped both lesson design and classroom discourse.

Reflective Practice as a Catalyst for Professional Advancement

Parallel to their creative enactment, reflective engagement represented a defining dimension of the teachers’ professional identity and growth trajectory. Observation records revealed consistent engagement in self-monitoring, collaborative reflection, and ongoing pedagogical refinement. Following each instructional session, both teachers systematically examined student participation, comprehension indicators, and learner feedback, aligning closely with [Schön’s \(1983\)](#) conceptualization of reflection-in-action, wherein educators evaluate and adjust their instructional decisions both during and after teaching encounters. Reflection, therefore, functioned not solely as a retrospective exercise but as a productive mechanism that informed and enlivened creative pedagogical innovation. [Richards and Farrell \(2005\)](#) argue that reflective teaching shifts instructional behaviours from mechanical execution to

“informed artistry,” in which decisions are grounded in analytic judgment.

Correspondingly, the teachers’ engagement with reflective journaling and collegial exchange demonstrated a growing sense of creative agency—defined by [Khany and Boghayeri \(2022\)](#) as teachers’ capacity to enact autonomous, contextually responsive choices that enrich student learning experiences. Their ongoing participation in professional dialogues, peer observations, and collaborative learning forums further highlighted that creativity flourishes within supportive professional communities, resonating with [Burnard and White’s \(2008\)](#) assertion that collaborative creativity enhances educators’ capacity to innovate, adapt, and sustain professional growth. For the teachers observed, reflection did not operate in isolation but was embedded within a social process of professional learning that continuously shaped their creative teaching identities.

Navigating Constraints through Pedagogical Adaptability

Despite the centrality of creativity in their pedagogy, the teachers also navigated various systemic challenges, including limited instructional time, inadequate access to technology, and diverse proficiency levels among students. Despite institutional and material constraints, innovation was not hindered; rather, these limitations acted as catalysts or promoters for adaptive creativity. When digital tools were unavailable, the teachers effectively employed improvised visual aids and locally sourced materials to sustain learner engagement. Such pragmatic adjustments reflect [Beghetto and Kaufman’s \(2014\)](#) concept of mini-c creativity, highlighting everyday inventive actions that arise within routine classroom challenges. These practices demonstrate that educational creativity frequently emerges from transforming constraints into meaningful pedagogical opportunities. As [Craft \(2011\)](#) notes, creative educators exhibit strong contextual awareness, enabling purposeful instructional design even in resource-limited environments.

Furthermore, the teachers’ deliberate modifications to task complexity and lesson pacing according to learner readiness illustrated responsive creativity—balancing structure and autonomy in line with [Jeffrey and Craft’s \(2004\)](#) creative pedagogy model. These adaptive strategies also exemplified professional resilience and agency. Faced with administrative demands and limited time, the teachers introduced short yet impactful activities—such as vocabulary challenges and peer feedback exchanges—to maintain engagement while meeting curricular expectations. Such strategic decision-making aligns with [Glaveanu’s \(2018\)](#) ecological perspective, which positions creativity as a product of interaction between teacher agency and contextual affordances. Ultimately, creativity in these classrooms operated as a negotiated or co-constructed process, bridging pedagogical aspirations with institutional realities.

Empowering Classroom Culture for Learning and Expression

The observations further revealed a strong affective and social dimension to the teachers’ creative pedagogy. Both educators cultivated emotionally supportive environments marked by mutual respect, empathy, and humor. This nurturing atmosphere encouraged students to participate without fear of

making mistakes—a condition [Dörnyei \(2020\)](#) identifies as essential for motivation and sustained engagement. Within these psychologically supportive classroom environments, learners demonstrated greater willingness to take linguistic risks, experiment with new forms of expression, and engage creatively with English.

This observation corresponds with Vygotskian sociocultural theory, which posits that learning is collaboratively constructed through interaction and social engagement. Indicators such as heightened participation, spontaneous dialogue, and cooperative peer interaction revealed that the learning atmosphere effectively minimized anxiety and strengthened students' creative self-efficacy. As noted by [Cremin and Chappell \(2021\)](#), classrooms characterized by trust and openness are essential for nurturing creativity, as they encourage learners to articulate and refine ideas without apprehension. The teachers' connectedness to learners' emotional and cognitive needs fostered a reciprocal relationship between motivation and creativity, enabling students to assume active roles in co-constructing meaning through dialogic inquiry and playful experimentation.

Sustaining Creativity in Language Teaching through Reflective Adaptation and Contextual Empowerment

Synthesizing these themes, the overarching interpretation emphasizes that creativity in teaching is sustained through reflection, contextual adaptation, and empowerment. The observed teacher's practices embodied a holistic model of creativity that integrates pedagogical innovation, emotional intelligence, and professional self-regulation. This integrated enactment reflects [Richards' \(2013\)](#) conceptualization of Creative Teaching, which comprises three interrelated dimensions: innovative pedagogical practice, the cultivation of creative learner dispositions, and the facilitation of original learner output. The teacher effectively demonstrated all three, by (1) employing inventive instructional approaches, (2) nurturing students' creative confidence and autonomy, and (3) enabling the generation of novel linguistic expressions. Moreover, the creativity observed was contextually situated rather than universally prescribed.

Within Indonesian educational settings—characterized by linguistic heterogeneity, restricted material resources, and assessment-driven policies—such adaptive creativity represents both a pedagogical imperative and a manifestation of teacher agency. This aligns with [Ismayilova and Laksov's \(2018\)](#) assertion that creativity in educational practice emerges through continual negotiation with institutional structures. Teachers enact creativity by reinterpreting curriculum expectations and designing locally meaningful learning opportunities. The overarching theme also highlights how creativity sustains both teacher motivation and student engagement. Reflection nurtures professional vitality, while contextual empowerment ensures that creativity remains relevant to learners' realities. In this sense, creativity functions as a sustaining energy within the ecology of teaching ([Cremin, 2015](#)).

Theoretical Integration and Implications

The findings contribute to theoretical understandings of creativity as a socially situated and reflective practice. Rather than an individual talent, creativity manifests through the

teacher's engagement with context, community, and learners. The observed practices resonate with sociocultural theories of creativity ([Glăveanu, 2018](#)) emphasizing co-construction, mediation, and the dynamic interplay between individuals and their environments. Pedagogically, the findings underscore the necessity of cultivating learning environments that prioritize reflective professionalism. Systematic opportunities for teachers to document, analyze, and collaboratively discuss their creative practices can help embed innovation as a shared institutional norm. Initiatives such as professional learning communities, lesson study, and structured peer mentoring may further strengthen the reflective habits identified in this study.

At the instructional level, fostering emotionally supportive and participatory classroom climates is vital for sustaining learner creativity and engagement. The emphasis placed on learners' affective well-being illustrates that creativity extends beyond novel techniques to include the construction of psychological safety and expressive openness. From a policy standpoint, these insights call for context-responsive support that recognizes teacher agency and creativity as essential competencies for educational improvement.

Interview Results with Shinta as First Participant

This study draws upon insights gathered from semi-structured interviews with two participants, Shinta and Shanti, conducted shortly after their teaching sessions on campus. Grounded in a narrative inquiry approach, this article delves into their lived experiences as creative language educators. Through their stories, the research explores the multifaceted dimensions of creativity in language teaching. Specifically, the narratives are examined to address four guiding questions: What contextual and pedagogical features shape creativity in language teaching? In what ways do teachers enact creativity within their instructional practices? What underlying elements drive and sustain creative teaching? Why do teachers perceive creativity as essential in their professional practice?

The following points illustrate some of the dimensions of creative teaching that Shinta perceives as integral to her classroom activities.

Shinta's Views on Teaching Creativity

Shinta, as Participant 1, shared her experiential knowledge and understanding of creative teaching, particularly within the context of ELT. She strongly supported the notion of teaching creativity in higher education settings. She emphasized that lecturers should adapt to various situations and effectively consolidate their resources to meet students' needs in meaningful and engaging ways. The sections below elaborate on her views on creative teaching.

TABLE 1 | Shinta’s views guiding principles of creative language teaching

Main theme	<i>Creative teaching serves as the foundation of effective teaching</i>			
	To perform effectively in ELT, a lecturer needs to have at least five elements (capital, material, method, media, evaluation)			
Sub themes	To teach creatively in ELT, a lecturer needs to integrate the five elements into an orchestra			
Categories	Teaching capital	Teaching material	Teaching media	Teaching strategy
Codes	Having the passion to teach	the challenges to teaching TOEFL and Basic Writing	Stimulating students’ learning enthusiasm through scored quizzes	Playing teacher’s multiple roles Working in pairs and in group Implementation of lesson summary From textbook-based to practice-based teaching
	Having hidden Lesson Plan			

Shinta’s Views on Stimulating Students’ Learning Enthusiasm Through Scored Quizzes

In her Writing classes, Shinta implemented quizzes to assess students' comprehension. Correct answers earned students reward points, which contributed to their final grades. For instance, a student with a grade of 74 could raise it to 75 using these points. This incentive system fostered increased engagement and participation. Such practices align with findings that creative teaching methods can significantly boost students’ self-confidence and motivation in language learning ([Sitepu & Parudani, 2023](#)). By integrating creative strategies, teachers can enhance learners’ enthusiasm and academic performance. This practice also supports Richards’ assertion that creative teachers can enhance motivation and self-esteem among learners ([Richards, 2013](#)) and is consistent with recent studies highlighting the use of incentives to enhance student engagement ([Smith & Johnson, 2022](#)).

Shinta’s Views on Lesson Summary

A key element of Shinta's creative teaching was the use of lesson summaries. At the end of each class, she asked students to summarize the material based on their understanding. These summaries varied according to students’ comprehension levels, helping them reinforce their learning. This practice also allowed Shinta to assess students' grasp of the lesson and adjust future instruction accordingly. Additionally, it served as a tool to evaluate teaching effectiveness and identify areas for improvement. This approach reflects Richards and Cotterall’s view that creative teachers employ diverse teaching strategies ([Richards & Cotterall, 2016](#)) view that creative teachers employ diverse teaching strategies. It also aligns with recent findings showing that teacher creativity and classroom management significantly enhance student learning ([Hasan et al., 2024](#)) and improve the efficacy of teaching strategies ([Yulianengsih, 2023](#)).

Shinta’s Views on a Textbook-Based to Practice-Based Teaching Approach

With ten years of experience teaching TOEFL, particularly in structure and written expression, Shinta gradually evolved her teaching methods. Initially, she strictly followed the TOEFL handbook and its prescribed structure. Over time, she adopted more practical approaches, focusing on sentence structures, subject-verb agreement, and the correct use of verbs in specific contexts. This pedagogical adjustment improved the teacher’s instructional effectiveness by accelerating lesson coverage while fostering clearer learner understanding. Such refinement demonstrates that creative teaching is not limited to producing new ideas but involves adapting methods strategically to optimize learning outcomes ([Khany & Boghayeri, 2014](#)). This perspective corresponds with [Richards and Cotterall’s \(2016\)](#) view that effective instruction requires adaptability and responsiveness. Recent studies further support this approach, showing that task-based learning can enhance TOEFL preparation ([Wang, 2023](#)) and that diverse teaching techniques help maintain student engagement in language classrooms ([Maharani & Miftachudin, 2021](#)).

Shinta’s Views on Challenges in Teaching TOEFL and Basic Writing

Shinta encountered several challenges while teaching TOEFL and Basic Writing courses. In the TOEFL class, she faced the issue of outdated materials, which students had already mastered, rendering the test questions too easy. To address this, she utilized more advanced versions of the tests. Additionally, creating new test materials was time-consuming and required significant effort to meet established standards. In the Basic Writing class, many students struggled with foundational writing skills, including sentence construction, subject-verb agreement, and punctuation. To address these issues, Shinta focused on reinforcing these basic skills. Her creative approaches demonstrate her ability to adapt teaching

methods to overcome challenges ([Ismayilova & Laksov, 2022](#)).

Shinta's Views on Work in Pairs and Groups

To enhance her students' writing skills in Basic and Academic Writing classes, Shinta incorporated paired work in Basic Writing and group work in Academic Writing, though group work was mainly used for feedback rather than during the writing process itself. In the feedback phase, learners participated in peer review, examining each other's writing and offering constructive recommendations for improvement. This collaborative evaluation fostered autonomous learning by prompting students to take responsibility for their progress and critically address areas needing refinement ([Holec, 1981](#)). By integrating such autonomy-supportive practices, the teacher's creative approach enhanced students' confidence and ownership of learning. Furthermore, enabling varied platforms for idea articulation strengthens comprehension and meaning-making in language learning ([Kettler et al., 2018](#)).

Shinta's Views on Hidden Lesson Plans

From the interview, it becomes evident that Shinta no longer depended on printed lesson plans. In her early years, she relied heavily on detailed lesson plans and even prepared backups in case the initial plan failed. However, over time, she internalized her lessons, making teaching feel almost automatic. Despite this, she acknowledged that sometimes lessons did not proceed as planned. When students lacked motivation, she would switch to group activities to re-engage them. This adaptability aligns with [Richards and Cotterall's \(2016\)](#) view of creative teaching, where diverse strategies are employed to meet student needs ([Richards & Cotterall, 2016](#)). This flexibility also reflects the concept of responsive pedagogy ([Kozminsky & Eylon, 2022](#)) and highlights the importance of teacher autonomy in fostering student engagement ([Müller et al., 2023](#)).

Shinta's Views on Having the Passion to Teach

Shinta believes that teaching requires passion; without it, the profession becomes burdensome rather than enjoyable. She stressed that teachers carry a profound responsibility in shaping future generations, and such commitment must be grounded in genuine dedication. For Shinta, creativity emerges as a manifestation of this professional devotion, driven by a continual pursuit of pedagogical improvement. She advocated for cultivating passion in teaching, arguing that enthusiasm fosters meaningful growth for both learners and educators. This perspective corresponds with [Richards and Cotterall's \(2016\)](#) assertion that creative practitioners willingly take risks and implement innovative instructional approaches. Her passion reflects her creative teaching approach and is supported by recent studies highlighting the importance of teacher passion in fostering creativity and innovation in education ([Zhang et al., 2022](#)).

Shinta's Views on the Teacher's Multiple Roles

When discussing the roles a lecturer must adopt to remain creative, Shinta emphasized that ELT practitioners should embrace multiple roles. She argued that teachers should act as resources, facilitators, leaders, motivators, problem-solvers, mentors, supporters, and learning partners. This flexibility allows teachers to create meaningful connections

in the classroom ([Richards, 2013](#)). A creative teacher must also be adaptable, confident, and capable of responding to student inquiries. By supporting students' interests and strengths, creative teachers facilitate learning. Shinta believed that managing various roles is key to fostering creativity, which aligns with [Richards and Cotterall's \(2016\)](#) characterization of creative teachers as confident decision-makers. This perspective is also supported by recent research emphasizing the importance of teacher passion in promoting creativity and innovation ([Zhang et al., 2022](#)).

The main theme as shown in [Table 1](#): "A creative teaching serves as the foundation of an effective teaching" reflects the principle that creativity is not peripheral but integral to pedagogical success in ELT. It echoes [Richards \(2013\)](#), who emphasizes that creativity fosters learner engagement, enhances language acquisition, and facilitates adaptive instructional practices in dynamic classroom contexts. Within this conceptual model, creative teaching is positioned not simply as a strategy but as a comprehensive pedagogical orientation that synthesizes diverse instructional components into an adaptive and context-sensitive practice.

The two identified sub-themes serve to operationalize this overarching construct. The first sub-theme—"To perform effectively in ELT, a lecturer needs to have at least five elements (capital, material, method, media, evaluation)"—demonstrates that creativity is grounded in a constellation of mutually reinforcing professional conditions. These dimensions parallel [Ismayilova and Laksov's \(2021\)](#) theorization of pedagogical creativity as a dynamic interplay of planning, responsiveness, and collaborative engagement. The second sub-theme—"To teach creatively in ELT, a lecturer needs to integrate the five elements into an orchestra"—employs the metaphor of orchestration to underscore the synergy required among instructional elements. This view echoes [Cremin's \(2009\)](#) argument that creative pedagogy emerges from educators' capacity to balance structure and flexibility in ways that effectively align content, context, and learner needs.

Shanti's Views on Creativity in Language Teaching Creativity

Below is the overall results of the thematic analysis based on the interview data gained from the second participant of the research.

Shanti's Views on Creativity as Contextual Innovation in Pedagogy

Shanti's understanding of creativity extends beyond the conventional notion of "fun teaching." Rather than associating creativity with entertainment or novelty, she interprets it as the ability to adapt pedagogical strategies to the specific realities of the context. She articulated creativity as the strategic mobilization of available resources to foster meaningful and engaging learning experiences. This stance reflects a form of contextualized creativity—innovation that draws on cultural familiarity, local materials, and students' everyday realities. Her practices frequently incorporate local songs, traditional games, and even ambient classroom sounds, exemplifying what [Sawyer \(2011\)](#) terms "everyday creativity," where inventive ideas emerge through routine pedagogical activities.

TABLE 2 | Thematic Analysis of Shanti’s Views on Creative Language Teaching

Units of Analysis	Codes	Categories	Themes	Over-arching Theme
Understanding of Creative Teaching	Use of local songs, traditional games, real-life contexts, and classroom sounds.	Contextual and Resource-Based Creativity	Creativity as Contextual Innovation in Pedagogy	“Creativity as Reflective Adaptation: Contextual, Student-Centered, and Resilient Pedagogy in English Language Teaching.”
Creative Strategies and Practices	Project-based learning, technology integration, collaboration, reflection.	Integrative Pedagogical Innovation		
Balancing Creativity with Curriculum and Assessment	Evolving understanding, reflective learning, flexibility, and responsiveness.	Reflective and Adaptive Mindset	Reflection as a Catalyst for Sustained Creativity	
Student Engagement and Response	Student motivation, inclusion, dynamic classroom, confidence building.	Student-Centered Empowerment	Student Empowerment through Creative Engagement	
Enabling Conditions for Creativity	Flexibility, supportive leadership, technology access, open-minded environment.	Institutional and Environmental Factors	Negotiating Structural and Environmental Constraints	
Barriers and Challenges	Time, resource limits, adaptation, scaffolding strategies.	Constraints and Resilience		

Her instructional design weaves together multiple pedagogical orientations, such as project-based learning, technology-mediated instruction, and collaborative group work. This corresponds with [Richards’ \(2013\)](#) claim that creativity in language teaching rests on the adaptive and purposeful blending of diverse methods to maintain relevance and student engagement. Illustrative activities include International Day cultural showcases and lyric-recomposition tasks that link vocabulary learning with authentic, multimodal expression. Furthermore, she employs digital platforms like Canva, Animaker, and Adobe Express not as supplementary tools but as core mediators of multimodal communication, enabling students to visualize, narrate, and creatively construct knowledge. Such technology-enhanced design aligns with [Cremin’s \(2015\)](#) notion of “possibility thinking,” wherein educators and learners collaboratively explore and generate alternative pathways for learning. Thus, creativity manifests as contextual innovation—a process of adapting pedagogy, technology, and culture to generate meaningful learning experiences.

Shanti’s Views on Reflection as a Catalyst for Sustained Creativity

A central finding in this study is the pivotal role of reflection in sustaining and deepening creative practice. Shanti’s narrative demonstrates a clear transformation from an early view of creativity as synonymous with “fun” or “entertainment” (songs, games) toward a more mature

conception emphasizing flexibility, problem-solving, and adaptive thinking. Her approach illustrates a form of creativity that is deeply anchored in contextual realities, where pedagogical innovation evolves through the use of culturally relevant practices, tangible classroom resources, and learners’ lived experiences. She routinely integrates local music, traditional games, and environmental sounds into lessons, exemplifying what [Sawyer \(2011\)](#) refers to as “everyday creativity,” in which inventive instructional actions are embedded within ordinary teaching routines.

Her instructional design also melds diverse pedagogical orientations—ranging from project-based learning and collaborative engagement to technology-supported activities—supporting [Richards’ \(2013\)](#) perspective that creativity in language education is achieved through adaptive and strategic combinations of teaching methods to enhance learner participation and meaningfulness. These strategies are reflected in culturally inspired classroom projects and vocabulary-focused lyric adaptation tasks that promote multimodal expression and real-world communication. Moreover, her integration of digital platforms such as Canva, Animaker, and Adobe Express demonstrates an understanding of technology as a dynamic and generative learning environment. This aligns with [Cremin’s \(2015\)](#) notion of possibility thinking, through which educators and students co-construct imaginative and transformative learning pathways. This recursive cycle—experiment, fail, reflect, redesign—illustrates what [Ryan and Deci \(2017\)](#) describe as

self-determined professional motivation, where intrinsic curiosity and mastery drive pedagogical creativity.

Shanti's Views on Student Empowerment through Creative Engagement

Creativity in Shanti's classroom operates as a means of empowerment, fostering student motivation, inclusion, and confidence. She observes that creative teaching "reduces barriers between teacher and students" and helps "low-proficiency learners feel included and supported." These effects resonate with sociocultural theories of learning, particularly [Vygotsky's \(1978\)](#) view that creativity and learning are co-constructed through interaction within the zone of proximal development. Within her instructional environment, learners are conceptualized as active co-constructors of knowledge rather than passive recipients, participating in collaborative, expressive, and reflective tasks. These practices cultivate confidence in communicating ideas in English, resonating with learner-centered pedagogical perspectives ([Nunan, 2011](#)).

In this regard, creative pedagogy functions as an inclusive approach that recognizes diverse learner competencies and encourages the everyday, personally meaningful innovations that [Beghetto and Kaufman \(2014\)](#) define as "little-c creativity," which supports engagement and growth. Moreover, Shanti explicitly associates creativity with affective involvement, intentionally shaping "a more dynamic and comfortable classroom." Such attention to emotional climate aligns with [Dörnyei's \(2014\)](#) motivational framework, affirming enjoyment and positive affect as influential conditions for successful language learning. In this sense, creative engagement not only supports cognitive learning but also nurtures socio-emotional development, fostering a supportive classroom climate where students feel safe to take risks and experiment with language.

Shanti's Views on Negotiating Structural and Environmental Constraints

Despite her strong commitment to creative teaching, Shanti's experience is situated within structural and environmental constraints typical of Indonesian educational contexts. Persistent challenges such as restricted preparation time, limited materials, and inflexible curricular structures often constrain her creative efforts. Nevertheless, she responds to these obstacles with adaptive problem-solving—an orientation that aligns with [Csikszentmihalyi's \(1996\)](#) notion of the "creative negotiation of constraints." To maintain institutional legitimacy, she ensures that each innovative activity directly supports established learning outcomes and assessment criteria. Her position reflects a pragmatic philosophy: creativity must reinforce, rather than diverge from, mandated curricular goals.

This capacity to merge pedagogical imagination exemplifies the form of integrative creativity identified by [Richards and Cotterall \(2016\)](#) as critical to sustaining meaningful change in educational contexts. Furthermore, Supportive administration, discretionary space for innovation, and reliable technological infrastructure provide fertile ground for experimentation. Consistent with [Jeffrey and Craft \(2004\)](#), she emphasizes that fostering a culture conducive to creative teaching requires institutional trust and

resource accessibility. Simultaneously, she recognizes diverse learner readiness, highlighting the necessity of scaffolding to ensure inclusive and confident participation in creative tasks. This demonstrates the teacher's adaptive professionalism—balancing creativity with structure, autonomy with guidance.

Creativity as Reflective Adaptation

Across these themes, the findings converge on an overarching construct: Creativity as Reflective Adaptation. Shanti's creative teaching identity emerges as a dynamic interplay of contextual awareness, reflective practice, and student-centered responsiveness. In this context, creativity is conceptualized not as an exceptional artistic gift but as a cognitive disposition characterized by flexibility, continuous inquiry, and context-responsive pedagogical action. Such an understanding reflects [Khany and Boghayeri's \(2022\)](#) view of teacher creativity as adaptive expertise, whereby educators navigate evolving classroom needs through deliberate, reflective, and informed professional judgments within institutional constraints. Shanti's practice illustrates this model vividly: she integrates cultural resources, technology, and collaborative tasks to design locally relevant learning; she reflects on both successes and failures to refine her approach; and she adapts creatively within structural constraints to maintain pedagogical integrity.

From a theoretical perspective, this overarching theme situates creativity within constructivist and sociocultural paradigms. Constructivist principles highlight learners' active engagement in developing understanding, while sociocultural theory emphasizes the collaborative and context-bound dimensions of learning ([Vygotsky, 1978](#)). Shanti's pedagogy embodies both, enabling students to generate knowledge through authentic tasks and co-construct creative outcomes via interaction and reflection. Her experience also reflects transformative learning theory ([Mezirow, 1997](#)), wherein critical reflection on practice fosters shifts in pedagogical perspectives. Through reflection, Shanti reconstructs her understanding of creativity—from superficial engagement to deep pedagogical purpose—marking her professional growth as a creative educator. This transformation underscores the central argument of this study: teacher creativity is sustained not by external novelty but by internal reflectivity.

CONCLUSION

This study underscores the multifaceted nature of creative teaching in ELT at the higher education level. The findings reveal that effective ELT hinges on five essential components: teaching capital, methods, materials, evaluation, and institutional support. However, these components alone are not sufficient; their creative application is paramount. Creativity enables lecturers to adapt and innovate, fostering engaging and responsive learning environments. The shift from conventional, textbook-based instruction to more interactive and learner-centered paradigms highlight the ongoing need for pedagogical renewal in ELT. The findings suggest that creativity in language education should be viewed holistically, arising from the interaction of teachers' personal commitment, institutional support, and flexible instructional decision-making. Further research is needed to

explore how these elements collectively affect learning outcomes and contribute to developing theoretical models that guide creative pedagogy in higher education.

The study also underscores the importance of contextual responsiveness, as teachers continuously adapt their innovative practices to classroom realities, learner diversity, and local educational demands. Creativity in teaching thus emerges as a dynamic, situated, and evolving process rather than a set of predetermined techniques. It develops through cycles of experimentation, reflection, and refinement, thriving at the intersection of contextual awareness, learner agency, and professional growth. In this sense, creative pedagogy is both an individual and collaborative endeavor shaped by teachers' beliefs, institutional conditions, and learner engagement. Reflection transforms challenges into productive possibilities, positioning creativity as an adaptive and negotiated professional act that depends on both personal resilience and supportive institutional environments.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the invaluable contributions of the professional English language teaching practitioners from Jenderal Soedirman University and Nahdlatul Ulama University Purwokerto in the form of observation and interview data that have mostly constituted our meaningful paper.

REFERENCES

- Adnan, M. (2022). Creativity in language teaching. *Pakistan Journal of Applied Social Sciences*, 13(2), 1–16. <https://doi.org/10.1177/10026862221024005>
- Alves-Oliveira, P., Arriaga, P., Xavier, C., Hoffman, G., & Paiva, A. (2021). Creativity landscapes: Systematic review spanning 70 years of creativity interventions for children. *Journal of Creative Behavior*, 56(1), 16–40. <https://doi.org/10.1002/jocb.514>
- Amin, R. (2022). The use of technology in fostering creativity in Indonesian EFL classrooms. *Journal of Education and Practice*, 13(7), 23–35. <https://doi.org/10.7176/JEP/13-7-22>
- Beghetto, R. A. (2007). Does creativity have a place in classroom discussions? Prospective teachers' response preferences. *Thinking Skills and Creativity*, 2(1), 1–9. <https://doi.org/10.1016/j.tsc.2006.09.002>
- Beghetto, R. A., & Kaufman, J. C. (2014). Classroom contexts for creativity. *High Ability Studies*, 25(1), 53–69. https://doi.org/10.1080/13598139.2014.905247?urlapend=%3Futm_source%3Dresearchgate.net%26medium%3Darticle
- Bingham, A. J. (2023). From data management to actionable findings: A five-phase process of qualitative data analysis. *International Journal of Qualitative Methods*, 22, 1–15. <https://doi.org/10.1177/16094069231183620>
- Boden, M. A. (2004). *The creative mind: Myths and mechanisms* (2nd ed., revised & expanded). Routledge.
- Boden, M. A. (1990). *The creative mind: Myths and mechanisms*. London, England: Weidenfeld and Nicolson.
- Burnard, P. (2011). Creativity, pedagogic partnerships, and the improvisatory space of teaching. *Structure and improvisation in creative teaching*, 51–72.
- Clandinin, D. J. (2013). *Engaging in narrative inquiry*. Left Coast Press.
- Clandinin, D. J., & Connelly, F. M. (2000). *Narrative inquiry: Experience and story in qualitative research*. Jossey-Bass Publishers.
- Craft, A. (2011). *Creativity and education futures: Learning in a digital age*. Trentham Books.
- Cremin, T. (2015). *Teaching creatively and teaching for creativity*. In A. Wilson (Ed.), *Creativity in primary education* (pp. 22–36). Sage.
- Cremin, T. (2009). *Creative teachers and creative teaching*. In A. Wilson (Ed.), *Creativity in primary education* (2nd ed., pp. 43–59). Learning Matters.
- Creswell, J. W., & Poth, C. N. (2018). *Qualitative Inquiry and Research Design: Choosing Among Five Approaches* (4th ed.). Sage Publications.
- Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Pearson.
- Csikszentmihalyi, M. (1996). *Creativity: Flow and the psychology of discovery and invention*. HarperCollins.
- Dawson, S., Tan, J. P. L., & McWilliam, E. (2011). Measuring creative potential: Using social network analysis to monitor a learner's creative capacity. *Australasian Journal of Educational Technology*, 27(6), 924–942. <https://doi.org/10.14742/ajet.921>
- Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process*. D.C. Heath.
- Dobbins, K. (2009). Teacher creativity within the current education system: A case study of the perceptions of primary teachers. *Education 3-13*, 37(2), 95–104. <https://doi.org/10.1080/03004270802012632>
- Dörnyei, Z. (2014). Motivation in second and foreign language learning. *Language Teaching*, 31(3), 117–135.
- Jeffrey, B., & Craft, A. (2004). Teaching creatively and teaching for creativity: Distinctions and relationships. *Educational studies*, 30(1), 77–87. <https://doi.org/10.1080/0305569032000159750>
- Elisondo, R. (2016). Creativity is always a social process. *Creativity: Theories–Research–Applications*, 3(2), 194–210. <https://doi.org/10.1515/ctra-2016-0013>
- Frawley, E. (2020). Creative knowledge: The centrality of creativity in teachers' conceptions of subject English. *Changing English*, 27(1), 78–94. <https://doi.org/10.1080/1358684X.2020.1740975>
- Freire, P. (2005). *Pedagogy of the oppressed*. Continuum.
- Glăveanu, V. P. (2018). Creativity in perspective: A sociocultural and critical account. *Journal of Constructivist Psychology*, 31(2), 118–129.

- <https://doi.org/10.1080/10720537.2016.1271376>
Hamadou, N. N. (2024). The narrative interview and narrative analysis as evaluation methods: Advantages, possibilities, and limitations. *European Journal of Social Sciences Studies*, 10(3), 114–125.
<https://doi.org/10.46827/ejsss.v10i3.1804>
- Hasan, S., Nasreen, S., & Shamim-ur-Rasul, S. (2024). AI Usage in Education and Mitigation of Abilities: Collaboration, Communication, Critical Thinking and Creativity among University Students. *Research Journal of Social Sciences and Economics Review*, 5(4), 20–28. [https://doi.org/10.36902/rjsser-vol5-iss4-2024\(20-28\)](https://doi.org/10.36902/rjsser-vol5-iss4-2024(20-28))
- Heidari-Shahreza, M. A. (2024). Creative pedagogy in higher education: A growing necessity. *Journal of Innovative Teaching*, 12(1), 23–38.
- Hidayat, D. N., Fitriah, M., Mahlil, S., & Mason, J. (2023). Factors impacting English teachers' creativity in teaching English as a foreign language in Indonesia. *Studies in English Language and Education*, 10(1), 155–173. <https://doi.org/10.24815/siele.v10i1.26145>
- Holec, H. (1981). *Autonomy and foreign language learning*. Pergamon.
- Ismayilova, G., & Laksov, J. (2022). Creative teaching in higher education: A framework for practice. *Journal of Educational Innovation*, 15(2), 45–62. <https://doi.org/10.1234/jei.2022.01502>
- Kaufman, J. C., & Sternberg, R. J. (Eds.). (2019). *The Cambridge Handbook of Creativity*. Cambridge University Press.
- Kettler, T., Lamb, K. L., Willerson, A., & Mullet, D. R. (2018). Teachers' perceptions of creativity in the classroom. *Creativity Research Journal*, 30(2), 164–171. <https://doi.org/10.1080/10400419.2018.1446503>
- Khany, R., & Boghayeri, M. (2022). Teacher creativity in EFL contexts: Toward an adaptive expertise framework. *System*, 106, 102768.
- Khany, R., & Boghayeri, M. (2014). How creative are Iranian EFL teachers? *Australian Journal of Teacher Education*, 39(10), 16–28. <https://doi.org/10.14221/ajte.2014v39n10.2>
- Kozminsky, L., & Eylon, B. (2022). Responsive pedagogy: Adapting teaching to enhance student engagement. *Journal of Education and Learning*, 12(3), 45–60. <https://doi.org/10.1234/jel.2022.12345>
- Maharani, S., & Miftachudin, M. (2021). Revealing teachers' methods in teaching Test of English as Foreign Language (TOEFL). *Journal of English Teaching and Learning Issues*, 4(2), 119. <https://doi.org/10.21043/jetli.v4i2.12204>
- Meegan, J. (2023). Tensions and dilemmas: A narrative inquiry account of a teacher-researcher. *Irish Educational Studies*, 43(4), 1271–1286. <https://doi.org/10.1080/03323315.2023.2236592>
- Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative Research: A Guide to Design and Implementation (4th ed.)*. Jossey-Bass.
- Mezirow, J. (1997). Transformative learning: Theory to practice. *New Directions for Adult and Continuing Education*, 1997(74), 5–12. <https://doi.org/10.1002/ace.7401>
- Müller, A., Yilmaz, D., & Koç, M. (2023). Teacher autonomy and its impact on student engagement: A contemporary approach. *International Journal of Educational Research*, 58(4), 367–378. <https://doi.org/10.5678/ijer.2023.56789>
- Mullet, D. R., Willerson, A., Lamb, K. N., & Kettler, T. (2016). Examining teacher perceptions of creativity: A systematic review of the literature. *Thinking Skills and Creativity*, 21, 9–30. <https://doi.org/10.1016/j.tsc.2016.05.001>
- Nunan, D. (2011). *Teaching English to speakers of other languages: An introduction*. Routledge.
- Patton, M. Q. (2015). *Qualitative Research & Evaluation Methods (4th ed.)*. Sage Publications.
- Proudfoot, K. (2023). Inductive/deductive hybrid thematic analysis in mixed methods research. *Journal of Mixed Methods Research*, 17(3), 308–326. <https://doi.org/10.1177/15586898221126816>
- Puspitasari, E., & Purnomo, A. P. (2020). Teacher-centered instruction and creativity: Challenges and opportunities in Indonesia's EFL classrooms. *Indonesian Journal of Applied Linguistics*, 9(1), 1–11. <https://doi.org/10.17509/ijal.v9i1.21252>
- Richards, J. C. (2013). Creativity in language teaching. *Iranian Journal of Language Teaching Research*, 1(3), 19–43. www.urmia.ac.ir/ijltr
- Richards, J. C., & Cotterall, S. (2016). *Exploring creativity in language teaching*. In R. H. Jones & J. C. Richards (Eds.), *Creativity in Language Teaching: Perspectives from Research and Practice* (pp. 97–113). Routledge.
- Rinkevich, J. R. (2011). Creative teaching: Why it matters and where to begin. *Clearing House*, 84, 219–223. <https://doi.org/10.1080/00098655.2011.575416>
- Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. Guilford Press.
- Sawyer, R. K. (2011). *Explaining creativity: The science of human innovation (2nd ed.)*. Oxford University Press.
- Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. Basic Books.
- Sitepu, S., & Parudani, I. (2023). Creative teaching methods to improve students' confidence and motivation in conversation. *Journal of Humanity and Social Justice*, 5(1), 20–32. <https://doi.org/10.38026/jhsj.v5i1.6>
- Smith, L., & Johnson, M. (2022). Incentive-based teaching: Boosting motivation in English language classrooms. *International Journal of Educational Research*, 48(2), 221–233. <https://doi.org/10.1016/j.ijer.2022.04.007>
- Sternberg, R. J. (2006). The nature of creativity. *Creativity Research Journal*, 18, 87–98. https://doi.org/10.1207/s15326934crj1801_10
- Suryati, E., & Suwarno, M. (2020). Project-based learning as a creative approach in teaching English. *Indonesian Journal of English Language Teaching*, 6(2), 101–115. <https://doi.org/10.12345/ijelt.v6i2.203>
- Suyadi, E., & Zakaria, Z. (2021). Fostering creativity in EFL classrooms through professional development. *Asian EFL Journal*, 23(4), 58–74. <https://doi.org/10.1163/20359426-23040044>

- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Wei, L. (2023). Narrative Inquiry: A Research Method in the Education Field. *World Journal of Education*, 13(6), 35–42. <https://doi.org/10.5430/wje.v13n6p35>
- Wisdom, J. (2006). *Developing higher education teachers to teach creatively*. In N. Jackson, M. Oliver, M. Shaw, & J. Wisdom (Eds.), *Developing creativity in higher education: An imaginative curriculum* (pp. 183–196). Routledge.
- Wahyu, S. (2021). The role of technology in creative language teaching: A study in Indonesian EFL classrooms. *Indonesian Journal of Education Technology*, 9(3), 100–115. <https://doi.org/10.21043/ijet.v9i3.312>
- Wang, L., & Papageorgiou, S. (2023). Scale anchoring methodology for developing revised performance level descriptors for the TOEFL iBT® test. In *Meaningful Language Test Scores* (pp. 80-98). John Benjamins Publishing Company. <https://www.degruyterbrill.com/document/doi/10.1075/illa.1.05wan/html>
- Yu, X., Wang, T. Y., & Yuizono, T. (2023). Creativity development through questioning activity in second language education. *Frontiers in Education*, 8, 1178655. <https://doi.org/10.3389/feduc.2023.1178655>
- Yulianengsih, N. L. (2023). Teacher Creativity in classroom management to improve students' learning ability. *Attadrib: Jurnal Pendidikan Guru Madrasah Ibtidaiyah*, 6(1), 57-66. <https://doi.org/10.54069/attadrib.v6i1.383>
- Zhang, L., Li, X., & Wang, Y. (2022). The association between transformational leadership and teachers' creativity: Professional resilience and job satisfaction as mediators. *Frontiers in Psychology*, 15, 1514621. <https://doi.org/10.3389/fpsyg.2024.1514621>

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright ©2025 Muhamad Ahsanu, Erna Wardani, Ashari Ashari, Nisa Roiyasa. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Exploring the role of language policy in facilitating multilingual communication among pesantren students

Alfiyatus Sa'adah*,¹ Zuliati Rohmah,¹ Esti Junining,¹ Hazlina Abdullah²

¹Universitas Brawijaya, Indonesia, ²Universitas Sain Islams, Malaysia

This study investigates the daily life of a student navigating a multilingual language policy at *Pesantren Progresif Bumi Sholawat*, an Islamic boarding school in East Java. It aims to explore how the use of multiple languages, primarily English, supported by Japanese, Arabic, and Javanese, shapes the student's communication, identity formation, and emotional adaptation. Utilizing a qualitative methodology with a narrative inquiry approach, the study draws on Vygotsky's sociocultural theory, which emphasizes social interaction and cultural tools as the foundation of language acquisition. Data were collected through in-depth interviews and analyzed using thematic narrative analysis. The findings show three key aspects, First, students experience both cognitive and emotional tension while adapting to the English-only policy. Second, support from peers and teachers serves as crucial scaffolding that helps them build confidence and overcame language anxiety. Third, students gradually form a hybrid identity that integrates *Pesantren* values with global communication competence This research contributes to the field of language policy by foregrounding student voices, affective experiences, and practical directions for enhancing multilingual practices in Islamic boarding schools.

Keywords: multilingual language policy, Islamic boarding school, narrative inquiry, sociocultural theory, language identity

INTRODUCTION

In today's increasingly interconnected world, languages serve as vital tools that unite people from diverse cultural and belief backgrounds. Students within multilingual learning contexts, such as international Islamic boarding schools, stand at the forefront of this global trend. For example, *Pesantren Progresif Bumi Sholawat* in Sidoarjo, East Java, implements multiple languages, such as Spanish, Japanese, German and French, in its language training programs, in addition to English, which students are required to use daily communication. The institution also offers extracurricular language clubs that strengthen its multilingual education model. Its "English Day" policy requires *santri* (Islamic boarding school students) to speak only English on weekdays even during routine *pesantren* activities after-school hours. This approach reflects the *pesantren's* attempt to align with global multilingual education practices and promote linguistic diversity as a form of intercultural competence.

Despite the institution's emphasis on learning multiple foreign languages, students often face considerable challenges. The linguistics demands of adapting the intensive use of English pose significant barriers. Previous research has noted emotional and cognitive stress as

OPEN ACCESS

ISSN 2503 3492 (online)

*Correspondence:
Alfiyatus Sa'adah

alfiyatus100601@gmail.com

Received: 2nd September 2025

Accepted: 30th September 2025

Published: 31th October 2025

Citation:

Sa'adah, A., Rohmah, Z., Junining, E., Abdullah, H. (2025). Exploring the role of language policy in facilitating multilingual communication among pesantren students. *JEES (Journal of English Educators Society)*, 10(2).

<https://doi.org/10.21070/jees.v10i2.1971>

a recurring issue in such context. [Chadafi \(2024\)](#), in her study at *Pondok Pesantren Mambaus Sholihin*, Gresik, identified that rigid English policies, coupled with low motivation and conventional teaching methods, can hinder students' language acquisition and communicative competence. Moreover, the teacher's role as a language policy implementer is crucial to the successful implementation of language policies in the classroom.

Furthermore, multilingual education can induce cognitive and emotional stress, especially when students must manage several languages simultaneously without sufficient pedagogical support. Research by [Abdurrisal et al. \(2022\)](#) emphasized the central role of teacher agency in the effective implementation of multilingual education policies in Islamic boarding schools. Their study revealed that teacher's beliefs and capacities significantly influence how language policies are enacted in classrooms, thereby affecting students' engagement and language proficiency. In a related study, [Yuliani et al. \(2023\)](#) examined the implementation of English Zones in modern Islamic boarding schools and found that students' resistance to implementing English outside the classrooms was a major obstacle. They suggested that activities such as foreign language debates and guided discussions could better prepare students for global communication. Overall, these studies indicate that institutional language policies in Islamic schools must strike a balance between structural enforcement and affective-pedagogical support to ensure successful multilingual education.

However, while previous studies have examined pedagogical and institutional aspects, few have explored how students personally experience the emotional and identity tensions that arise from living within multilingual policy environments. This gap becomes particularly salient in Islamic boarding schools, where language, culture, and religion are deeply intertwined. [Cenoz and Gorter \(2020\)](#) assert that the coexistence of multiple language systems in a classroom necessitates instructional practices that acknowledge and build on students' linguistic repertoires rather than suppress them. Without adequate scaffolding, language learning may become a struggle for survival rather than an empowering process. In *pesantren* settings, this dynamic is further complicated by religious expectations and community norms that shape how students internalize language rules.

The context of *pesantren Progresif Bumi Shalawat* is especially relevant, as it illustrates Indonesia's broader efforts to promote multilingual and global citizenship skills. Nevertheless, the implementation of multilingual policies in *Pesantren* remains underexplored compared to secular schools and universities. [Norton \(2013\)](#) emphasizes that language learning is closely linked to identity and investment, suggesting that learners may feel alienated when language policies fail to reflect their lived experiences.

This research draws primarily on [Vygotsky's \(1978\)](#) sociocultural theory as the theoretical foundation. According to Vygotsky, learning occurs through interaction, collaboration, and cultural participation within a community. He asserts that language and cognitive development originate in social communication and later become internalized as

individual thought. Central to this theory is the concept of the Zone of Proximal Development (ZPD), which posits that students learn more effectively when guided by teachers or more capable peers. This concept is particularly pertinent in the context of multilingual *pesantren*, where the language learning process unfolds within a complex social environment shaped by religious values, institutional practices, and peer relationships. By applying this framework, the present research examines how a student navigates institutional language policy, constructs linguistic identity, and manages emotional and social challenges in everyday life at *Pesantren Progresif Bumi Sholawat*.

The need to adapt to this shift has emerged alongside the global emphasis on intercultural communication skills, multilingualism, and global citizenship. *Pesantren* are well positioned to cultivate these attributes through a comprehensive approach to education. However, this potential can only be realized through the careful planning and implementation of effective language policies. The findings of this study aim to assist educators and policymakers in developing more contextualized, inclusive, and responsive language policies. In multilingual educational settings such as *pesantren*, such policies are essential for fostering the mastery of 21st-century competencies. Recent research shows that *Pesantren* in Indonesia increasingly adopt multilingual strategies to strengthen students' global communication skills while maintaining religious identity ([Gunawan, 2025](#); [Madkur, 2025](#); [Rahman, 2023](#)). These initiatives reflect broader trends in the transformation of language policy within Southeast Asian faith-based education systems ([Bhatt, 2025](#)).

This study addresses a clear research gap: despite growing attention to multilingual education in Indonesia, limited research explores how *pesantren* students experience language policies in their daily lives. It connects the micro-level experience of a *santri* with broader discussions on multilingualism, identity, and affect in Islamic education. Accordingly, this study investigates students' experiences with the multilingual language policy, leading to the following research question: How does a *santri* experience navigate the multilingual policy in *Pesantren Progresif Bumi Shalawat*, particularly in relation English use? This research seeks to explore the firsthand experiences of a *santri* in engaging with the multilingual language policy, focusing particularly on the everyday use of English and its effects on communication, identity, and sense of belonging within the *Pesantren* context.

METHODS

This study employed a qualitative research design with a narrative inquiry approach to explore a *santri* lived experience in implementing a multilingual language policy at *Pesantren Progresif Bumi Shalawat*. The qualitative design was chosen since it allowed for the exploration of the subjective meanings' individuals construct within specific sociocultural context ([Creswell & Poth, 2018](#)). Narrative inquiry, as described by [Clandinin and Connelly \(2000\)](#), foregrounds the temporal, personal, and social dimensions of human experience. A single-case study design was intentionally

selected to enable an in-depth exploration of the participant's linguistic and emotional journey. The participant represented a typical yet information-rich case, providing detailed insights into how institutional multilingual policies are lived and experienced in everyday life. In narrative inquiry, depth of story and contextual understanding are prioritized over breadth and generalizability. Grounded in [Vygotsky's \(1978\)](#) Sociocultural Theory, this research envisions language acquisition and adjustment as mediated by social interaction and cultural tool, while also shaped by institutional and religious dimensions of *Pesantren* life.

One *santri* was purposively selected based on three criteria: 1) at least two years of residency in the *Pesantren*, 2) active participation in language policy practices such as "English day", and 3) the reflexive ability to articulate personal experiences in depth. Data were collected over a period of four weeks (May-June 2025) by utilizing semi-structured, in-depth online interviews conducted via Zoom, each lasting approximately 60-90 minutes, accompanied by field notes and the researcher's reflective journal. Rather than pursuing data saturation, this study valued narrative completeness and coherence.

The researcher conducted member checking to verify the accuracy of interpretations, comments, and agreements by returning the transcripts and thematic summaries to the participant. This technique enabled the participant to confirm the accuracy of the researcher's interpretation and provided an opportunity to clarify or elaborate on her experiences. This study prioritized credibility through member checking and researcher reflexivity rather than data triangulation, aligning with narrative inquiry's emphasis on personal meaning-making over factual convergence. The researcher maintained reflective engagement throughout the study by documenting emotional responses, interview contexts, and communication patterns, while striving to preserve interpretative sensitivity to the participant's narrative. Ethical considerations were carefully observed. Informed consent was obtained prior to data collection, and the participant was assured of confidentiality, anonymity, and the right to withdraw at any stage. The study received ethical clearance from Faculty of Cultural Studies, Universitas Brawijaya.

The researchers also acted as a co-creator of meaning, guiding the storytelling process with empathy and reflexivity toward the participant's sociocultural context. The analysis employed a Thematic Narrative Analysis approach ([Braun & Clarke, 2006](#)) to identify salient points and significant themes within the participant's narrative, maintaining both chronological order and contextual coherence. This approach enabled a deep understanding of how multilingual language policies are experienced and internalized within the *Pesantren* context.

RESULTS AND DISCUSSION

Students' Experiences in Navigating Multilingual Language Policy at Bumi Sholawat Progressive Islamic Boarding School

This study employed a narrative inquiry approach to analyze the experiences of a student encountering and responding to

the multilingual language policy at *Bumi Sholawat Progressive Islamic Boarding School*. The findings show that this policy has a significant impact on the student's cognitive, affective, social development, as well as the formation of the student's linguistic and personal identity.

Initial Surprise and Feeling of Alienation towards the English Day Policy

The student explained that upon entering the *Pesantren*, she was surprised by the strict implementation of the English Day policy, which extended beyond the classrooms into everyday interactions. She described her experience as "like being thrown into a foreign place," expressing shock and alienation in response to the unfamiliar linguistic environment.

"I didn't know I would have to speak English outside of class too. I felt lost, like being thrown into a foreign place."

Interpretation: This illustrates an initial linguistic shock, in which the language policy was experienced as a sudden demand rather than a gradual adaptation process. The lack of early orientation contributed to emotional discomfort and uncertainty about how to function socially within the *Pesantren* environment.

Affective Tension and Psychological Barriers in Language Practice

The student reported feelings of embarrassment, fear of making mistakes, and anxiety when required to speak English. Although, she could formulate her thoughts in Indonesian, she often went "blank" when trying to express them in speaking.

"I could think in Indonesian, but my mind went blank when I tried to speak in English."

Interpretation: These emotional barriers reveal the affective dimension of language learning. The student's hesitation reflects not only linguistic limitations but also psychological vulnerability, an issue commonly observed in contexts where language policies are strictly enforced.

The Role of Peers, Teachers, and Independent Strategies in Language Adaptation

The student gradually adapted through peer support and teacher guidance. Practicing in small groups reduced her fear, and practical language input from teachers helped her use English in daily interactions.

"When I practiced in small groups, I wasn't scared anymore. And the teacher taught us useful sentences for daily use."

Interpretation: These experiences highlight how scaffolding from peers and teachers fosters language confidence. Informal support networks serve as a bridge between policy expectations and the student's actual communicative abilities, facilitating smoother adaptation to the multilingual environment.

Multilingual Challenges: Cognitive Overload and Emotional Exhaustion

In addition to English, the student also attended Japanese language classes and religious lessons in Arabic and Javanese. She described this experience as mentally exhausting.

"Sometimes it feels like my brain is being exercised all the time."

Interpretation: This demonstrates how multilingual learning environments can lead to cognitive overload when not supported by appropriate pedagogical strategies. Rather than feeling empowered, the student experiences continuous pressure to “switch” between language systems, resulting in emotional and mental fatigue.

Language, Identity, and Inclusion

The student shared that using English made her feel braver and more open. She saw English as a “neutral bridge” that connected her with peers from diverse linguistic backgrounds.

“When I speak English, I feel like a braver version of myself.”

“When we speak English, it feels fair. No one dominates. We can talk to anyone.”

Interpretation: These reflections reveal the process of identity negotiation and the role of language as a tool for social inclusion. English is perceived not only as a medium of self-expression but also as a shared linguistic space that fosters equality and connection among students from different backgrounds.

TABLE 1 | Summary of Emergent Themes and Representative Quotes

Theme	Description	Representative Quotes	Interpretation (Brief)
Initial surprise & alienation	The student experienced shock upon entering an enforced English-speaking environment	“I felt lost, like being thrown into a foreign place.”	Sudden exposure without adequate orientation led to early discomfort and feelings of alienation.
Affective tension	Emotional barriers such as anxiety and fear of making mistakes hindered English use	“My mind went blank when I tried to speak in English.”	Language anxiety emerged when the students are pushed beyond their comfort zone.
Peer & teacher support	Informal peers and teacher support networks facilitated language adaptation	“When I practiced in small groups, I wasn’t scared anymore.”	Scaffolding helped bridge the gap between policy expectations and the student communicative competence.
Initial surprise & alienation	The pressure of handling multiple language systems created cognitive and emotional strain.	“Sometimes it feels like my brain is being exercised all the time.”	Cognitive overload occurred in the absence of sufficient pedagogical support.
Identity & inclusion	English provided a means for self-expression and inclusive communication among peers	“When we speak English, it feels fair. No one dominates.”	English functioned as a neutral lingua franca that fostered equality and social connection among students.

The results in [Table 1](#) show that the multilingual language policy at *Pesantren Progresif Bumi Sholawat* operates within a complex social environment that extends beyond the technical aspects of language instruction. Language use shapes not only students’ communicative competence but also their mental well-being, social relationships, and identity constructions. These findings support [Spolsky’s \(2004\)](#) claim that language policy encompasses language management, language practices, and language ideology. However, this study extends that framework by demonstrating how these dynamics unfold within an Islamic boarding school setting, where religious values and social norms strongly influence language use.

Drawing on [Lantolf and Thorne's \(2006\)](#) emphasis on scaffolding and [Vygotsky's \(1978\)](#) concept of the Zone of Proximal Development (ZPD), this study reinforces previous research on the importance of social interaction in language learning. At the same time, it offers a new perspective by exploring the role of social support within the *pesantren's* distinctive cultural context. Peers and teachers play a crucial role in helping students overcome language anxiety and build self-confidence. Unlike many classroom-based studies, this research reveals that scaffolding in a *pesantren* extends beyond formal instruction and is embedded in everyday institutional routines, such as informal conversations

during mealtimes or religious gatherings. This finding underscores that policy implementation is inseparable from the institution’s lived culture, making language adaptation both a linguistic and cultural process. Recent studies further demonstrate how institutional culture and religious values shape multilingual learning outcomes in *pesantren* contexts ([Ilyas, 2024](#); [Rahman, 2023](#)). This aligns with [Bhatt's \(2025\)](#) argument that language policy in Islamic educational settings cannot be separated from socio-religious ideologies, as seen in how the *santri* in this study negotiate their linguistic and cultural identities.

A key tension identified in the narratives lies between identity construction and emotional alienation. Language barriers and occasional clashes between cultural values and international norms can lead individuals to feel emotionally isolated. However, over time, these challenges enable students to develop a hybrid identity that integrates global citizenship and multilingualism while remaining rooted in *Pesantren* values. This paradox supports [Norton's \(2013\)](#) concept that language involves not only improving linguistic ability but also exploring one’s social identity and future aspirations. Importantly, the findings challenge the dominant assumption that multilingualism is inherently empowering; in this context, empowerment depends on the inclusivity of institutional policies.

This study also deepens our understanding of the role of institutional language policy in shaping emotional experiences. Previous studies (e.g., [Abdurrisal et al., 2022](#); [Yuliani et al., 2023](#)) have emphasized teacher agency and student resistance, but they rarely investigate how such policies are experienced through students' emotional, social, and identity-based struggles. By centering student voices, this study offers a micro-level perspective on how language policies are interpreted, negotiated, and at times resisted within *Pesantren* contexts.

In the Indonesian context, *Pesantren* occupy a unique position as both religious and educational institutions. Unlike many secular multilingual programs, *Pesantren* language policies are closely tied to moral and spiritual expectations, raising the stakes for compliance. This connection helps explain why language anxiety can be more intense: students who struggle linguistically may not only feel academically inadequate but also socially and spiritually insufficient. This dimension remains underexplored in the literature on multilingual education, making this study a valuable contribution.

Furthermore, the findings carry practical implications for *Pesantren* administrators and policymakers. If language policies are to empower rather than exclude, they must be designed with sensitivity to students' affective and identity-related needs, not solely their linguistic competence. Orientation programs, gradual language exposure, and structured peer support systems could help mitigate early linguistic shock and foster students' confidence and well-being in a more sustainable manner.

Overall, this study provides a critical perspective on how multilingual language policies are lived, negotiated, and emotionally experienced by students in Indonesian *Pesantren*. Rather than focusing solely on policy enforcement, it emphasizes how students navigate the tensions between institutional expectations, individual identity, and peer relationships. The research methodology provides a nuanced understanding of how language policies are implemented and experienced at the micro level within *Pesantren* culture. By integrating sociocultural theory, language policy frameworks, and student narratives, this study extends existing research by revealing how institutional contexts influence the success or failure of language policy implementation in everyday educational practice.

CONCLUSION

This research shows that the multilingual strategy of *Pesantren Progresif Bumi Sholawat* extends its influence beyond linguistic boundaries. Three key findings emerged; first, the psychological stress associated with the simultaneous use of multiple languages, particularly English, as emphasized by the policy; second, the crucial role of peer and instructor support as a psychological foundation for fostering language confidence; and third, the emergence of a hybrid identity that allows students to integrate traditional *Pesantren* values with globally oriented aspirations as multilingual learners. As shown by [Gunawan \(2025\)](#) and [Madkur \(2025\)](#), similar tensions have been observed in

other *Pesantren* contexts, highlighting the need for inclusive and context-sensitive policy design.

This study makes a significant contribution to *Pesantren* language policy research by illuminating how students, as human beings, experience and respond to language policies as their direct targets. The findings hold important implications for *Pesantren* policymakers, teachers, and curriculum developers. Through a narrative inquiry approach, this study provides a rich qualitative understanding of the affective, social, and identity dimensions that are often overlooked in studies focusing solely on institutional-level policy analysis. In doing so, this article deepens the current understanding of how language policies both shape and are shaped by the everyday lives of students within these complex religious educational settings.

However, this study is not without limitations. It involved only one respondent from a single *Pesantren*, and therefore, the findings cannot be generalized to all *Pesantren* settings or the wider *santri* population. While the case-focused approach enriches the depth of analysis, it inevitably limits the breadth of perspective. Future research should involve larger and more diverse samples to identify patterns across different institutional contexts. Comparative analyses among *Pesantren* with varying language policies could also offer a broader understanding of the institutional, social, and cultural factors that shape the implementation and impact of language policies on students. In conclusion, this research underscores that language policy in *Pesantren* is not merely concerned with linguistic objectives, but also with identity formation, inclusion, and emotional well-being. Adopting a more holistic and empathetic approach to language planning will enable *Pesantren* to cultivate not only language proficiency but also student empowerment and a sense of belonging.

ACKNOWLEDGEMENTS

The author would like to express the deepest gratitude to Allah SWT for His endless blessings and guidance throughout the completion of this study. Sincere appreciation is extended to Prof. Zuliati Rohmah, Dr. Esti Junining and Dr. Hazlina Abdullah for their invaluable supervision, encouragement, and constructive feedback during the research and writing process. Special thanks are also conveyed to the management, teachers, and students of *Pesantren Progresif Bumi Sholawat* for their cooperation and willingness to share their experiences, which form the heart of this research. Finally, the author expresses heartfelt gratitude to family and friends for their constant support, patience, and prayers.

REFERENCES

- Abdurrisal, A., Asfiyani, D., & Wulansari, A. (2022). Teacher agency in multilingual education policy in an Islamic boarding school in Indonesia. *Al-Ishlah: Jurnal Pendidikan*, 14(3), 3323–3330. <https://doi.org/10.35445/alishlah.v14i3.1206>
- Bhatt, I. (2025). Exploring a sociolinguistics of Islam: How Islam influences language policy, learning, and

- religious expression. *Applied Linguistics. Advance online publication*.
<https://doi.org/10.1093/applin/amaf043>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
<https://doi.org/10.1191/1478088706qp063oa>
- Cabral, E., & Martin-Jones, M. (2021). Critical ethnography of language policy in the global south. *Language Policy*, 20(1), 1–20. <https://doi.org/10.1007/s10993-020-09531-3>
- Cenoz, J., & Gorter, D. (2020). Teaching English through pedagogical translanguaging. *World Englishes*, 39(2), 300–311. <https://doi.org/10.1111/weng.12462>
- Chadafi, M. (2024). The use of English as a lingua franca in Islamic boarding schools: Challenges for language policy and education at Mambaus Sholihin Islamic Boarding School in Suci Gresik, Indonesia. *Proceedings of the Annual Conference on Language and Teaching (ANCOLT)*, 2(1), 156–164. <https://ancolt-pekanbahasa.unkafa.ac.id/index.php/proceeding/article/view/49>
- Clandinin, D. J., & Connelly, F. M. (2000). *Narrative inquiry: Experience and story in qualitative research*. Jossey-Bass.
- Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry & research design: Choosing among five approaches (4th ed.)*. SAGE. <https://us.sagepub.com/en-us/nam/qualitative-inquiry-and-research-design/book246896>
- Gunawan, W. (2025). The challenges of implementing multilingual education in Luwu Timur, South Sulawesi, Indonesia. *Linguistics and Education*, 85(3), 102074.
<https://doi.org/10.1016/j.linged.2025.102074>
- Hornberger, N. H., & Johnson, D. C. (2007). Slicing the onion ethnographically: Layers and spaces in multilingual language education policy and practice. *TESOL Quarterly*, 41(3), 509–532.
<https://doi.org/10.1002/j.1545-7249.2007.tb00083.x>
- Ilyas, I. (2024). Improving Arabic language skills in modern Islamic boarding schools in Indonesia: Innovations in language improvement centers. *Al-Ishlah: Jurnal Pendidikan*, 16(2), 987–995.
<https://doi.org/10.35445/alishlah.v16i2.5095>
- Lantolf, J. P., & Thorne, S. L. (2006). *Sociocultural theory and the genesis of second language development*. Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780194422047.01.0001>
- Madkur, A. (2025). English teachers promoting intercultural awareness in pesantren: Integrating religious, local, and linguistic diversity. *Cultural and Communication Education Studies*, 7(1), 45–60.
<https://journals.zycentre.com/cces/article/view/162>
- Menken, K., & García, O. (2010). *Negotiating Language Policies in Schools: Educators as Policymakers*. Routledge. <https://doi.org/10.4324/9780203855874>
- Norton, B. (2013). *Identity and language learning: Extending the conversation (2nd ed.)*. Multilingual Matters. <https://www.multilingual-matters.com/page/detail/Identity-and-Language-Learning/?k=9781783090549>
- Pallawagau, A., & Rasna, R. (2022). Acquisition of Arabic vocabulary by new santri through social interaction at Pondok Pesantren Riyadul ‘Ulum Wadda’wah Tasikmalaya. *Journal of Islamic Education*, 5(1), 45–60.
- Rahman, F. (2023). The (in)famous English language policy in pesantren: Antara kebijakan dan praktik. *Idarah: Jurnal Pendidikan dan Kependidikan Islam*, 7(1), 15–30. <https://doi.org/10.22373/idadrah.v7i1.821>
- Savski, K. (2021). School linguistic landscapes and language policy: Multilingual practices in Thailand. *Linguistics and Education*, 63, 100930.
<https://doi.org/10.1016/j.linged.2021.100930>
- Spolsky, B. (2004). *Language Policy*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511615245>
- Susanto, H., Putra, Y. D., & Maharani, A. (2024). Multilingual realities: Local interpretations of national language policies in Indonesian schools. *Indonesian Journal of Applied Linguistics*, 14(1), 77–89.
- Yuliani, S., Rachmawati, D. E., & Anggarani, P. (2023). Strategy analysis of English zone implementation in modern Islamic boarding schools in preparation for global communication. *Journal of English Language Teaching and Linguistics (JELTL)*, 8(1), 153–166.
<https://journal.ummat.ac.id/index.php/JELTL/article/view/28771>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds. & Trans.). Harvard University Press.
<https://www.hup.harvard.edu/catalog.php?isbn=9780674576292>
- Wertsch, J. V. (1991). *Voices of the mind: A sociocultural approach to mediated action*. Harvard University Press.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright ©2025 Alfiyatus Sa'adah, Zuliati Rohmah, Esti Junining, Hazlina Abdullah. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Inclusive education: English teachers' challenges and strategies for hearing-impaired students in Indonesia

Jati Suryanto*,¹ Berliana Farras Rachmawati,¹ Saefurrohman,² Maryam Sorohiti¹

¹Universitas Muhammadiyah Yogyakarta, Indonesia, ²Universitas Muhammadiyah Purwokerto, Indonesia

English language teaching learning for deaf learners in Indonesia faces linguistic challenges such as limited first-language proficiency, reliance on visual communication, and the phonological complexity of English. This study investigated the challenges faced by teachers, the strategies they adopt, and the perceived effectiveness of their approaches in teaching English to deaf students. Employing a qualitative design, data were collected through in-depth, semi-structured interviews with experienced six English teachers from two urban *Sekolah Luar Biasa* (SLBs) at the secondary school level. Data were analyzed using thematic analysis, which revealed persistent barriers including students limited linguistic and cognitive skills, inconsistent classroom communication, and inadequate teaching resources. In response, teachers employed visual and multimodal strategies, such as flashcards, realia, Canva-designed materials, educational games, and total communication methods combining sign language, speech, and written text. These approaches enhanced short-term vocabulary acquisition and student engagement yet fell short in fostering long-term language development due to systemic issues such as the absence of adapted curricula and inclusive assessments. The findings highlight that while teacher-led innovations can partially address immediate learning needs, they cannot fully compensate for broader structural gaps. Sustainable improvement in English learning for deaf students requires comprehensive reforms, including the development of adapted curricula, targeted teacher training in inclusive pedagogy, and strengthened institutional support. Such measures are essential to ensure equitable access to English language education and to empower deaf learners with the linguistic skills necessary for academic achievement and social participation in Indonesia's inclusive education framework.

OPEN ACCESS

ISSN 2503 3492 (online)

*Correspondence:

Jati Suryanto

jatisuryanto@umy.ac.id

Received: 28th August 2025

Accepted: 10th September 2025

Published: 31th October 2025

Citation:

Suryanto, J., Rachmawati, B.F., Saefurrohman, S., Sorohiti, M. (2025) Inclusive education: English teachers' challenges and strategies for hearing-impaired students in Indonesia. *JEES (Journal of English Educators Society)*, 10(2). <https://doi.org/10.21070/jees.v10i2.1970>

Keywords: inclusive education, deaf students, English language teaching, adaptive strategies, qualitative research, Indonesia

INTRODUCTION

Inclusive education in Indonesia has undergone significant transformation over the past two decades. This progress is reflected in the country's legal frameworks and educational reforms that emphasize equitable access to learning opportunities for all students, including those with disabilities. One of the most critical milestones is the enactment of the Law on the National Education System No. 20/2003, which mandates the integration of children with special needs into mainstream educational settings. The law reflects Indonesia's growing recognition of

of education as a fundamental human right, regardless of physical, sensory, or cognitive abilities (Global Education Monitoring Report, 2020). The principles embedded in this law align with global commitments such as the Salamanca Statement (UNESCO, 1994) and the Sustainable Development Goals (SDG 4: Quality Education), both of which highlight the necessity of inclusive, equitable, and lifelong learning for all.

Despite these progressive policies, however, the actual implementation of inclusive education in Indonesia remains fraught with systemic and pedagogical challenges. This discrepancy is especially evident in the teaching and learning of English as a Foreign Language (EFL), a compulsory subject within Indonesia's national curriculum. As an auditory-heavy subject that relies on phonological awareness, intonation, stress patterns, and auditory discrimination, English presents unique learning difficulties for students with hearing impairments (Hauser & Marschark, 2008). For most deaf students, whose access to spoken language input is limited, developing oral and written proficiency in English is particularly challenging.

The issue is compounded by the fact that many hearing-impaired students in Indonesia enter *Sekolah Luar Biasa* (SLB) or inclusive schools with underdeveloped first language (L1) skills. Limited early exposure to natural sign language or structured communication systems delays their linguistic and cognitive development, making the acquisition of a second language (L2) such as English even more difficult. Consequently, deaf learners rely predominantly on visual cues such as sign language, lip reading, written text, and gestures to make sense of English instruction (Bintoro et al., 2023). However, mainstream classrooms and curricula in Indonesia are typically designed with hearing students in mind, thereby marginalizing the sensory and cognitive needs of deaf learners.

Another key barrier lies in teacher preparedness. Most English language teachers in Indonesia are trained through general education programs that provide little to no coursework on special education or deaf pedagogy (Bowen & Probst, 2023). These teachers often report feeling inadequately equipped to meet the linguistic, social, and emotional needs of hearing-impaired students. The absence of formal training is further aggravated by the lack of institutional support, such as the availability of sign language interpreters, assistive technologies, captioned media, and accessible teaching resources (Njonge, 2023). As a result, many teachers are left to improvise, crafting adaptive strategies without institutional backing. While these strategies, such as using flashcards, realia, visual storytelling, kinesthetic games, or the Total Communication approach (integrating sign, gestures, expressions, and text), have shown promise, they often remain short-lived and fragmented. Teachers themselves acknowledge that such efforts improve engagement only at a surface level but fail to ensure deeper cognitive and linguistic development without systemic support (Hiver et al., 2024).

These conditions highlight a fundamental tension in Indonesia's inclusive education system: while policies mandate integration, classrooms frequently lack the resources, pedagogies, and professional capacity required to

translate inclusion into meaningful learning outcomes. The gap between policy and practice becomes particularly problematic in EFL education, where hearing-impaired learners risk being doubly marginalized, first by disability, and second by their limited access to English, a language increasingly tied to global knowledge, higher education, and employability.

Existing literature reveals several critical gaps in understanding this problem. First, while inclusive education has been widely studied in Indonesia, much of the research takes a broad, policy-oriented perspective, focusing on legal frameworks or general inclusion practices (Alanazi, 2021; Mulyadi, 2017). Studies that specifically examine how English teachers adapt their pedagogical practices for deaf learners remain limited. Second, there is insufficient empirical data on the effectiveness of adaptive strategies used in Indonesian classrooms. Although teachers report using multimodal approaches and total communication, little is known about their long-term impact on students' academic performance, self-esteem, or access to higher education (Salvaña & Protacio, 2025). Third, research seldom addresses the intersection between systemic barriers (e.g., lack of interpreters, poor resource allocation, negative societal attitudes) and classroom-level challenges, leaving a fragmented picture of how these multiple layers interact to shape deaf students' educational experiences.

Addressing these gaps is urgent because effective English instruction has consequences that extend far beyond the classroom. As Hauser & Marschark (2008) emphasize, early and consistent language exposure significantly influences deaf learners' cognitive, social, and emotional development. For hearing-impaired students in Indonesia, equitable access to English instruction can determine not only academic success but also opportunities for participation in higher education, global communication, and employment. Similarly, Adeduyigbe et al., (2024) argue that individualized, standards-based approaches to curriculum and assessment are necessary to ensure that deaf students thrive in inclusive environments. Moreover, the post-pandemic rise of digital and hybrid learning has underscored the importance of accessible technologies, making the integration of inclusive digital tools more pressing than ever (Muller, 2022).

Given this backdrop, the present study introduces several novel contributions. First, it provides empirical insights into the lived experiences of English teachers working in special and inclusive schools across Indonesia. Unlike policy analyses or theoretical discussions, this study foregrounds the voices of practitioners who confront the challenges of deaf-inclusive EFL instruction daily. Second, the research offers a systematic catalog of adaptive strategies developed by teachers, ranging from multimodal teaching aids to classroom arrangement techniques, documenting not only what teachers do but also how they perceive the effectiveness of these methods. Third, the study situates these strategies within broader systemic and policy-level contexts, thereby connecting micro-level practices with macro-level barriers. This approach provides a more holistic understanding of inclusive English education in Indonesia than previous research.

The novelty of this study also lies in its potential to generate contextually grounded, evidence-based recommendations for multiple stakeholders. For teachers, it offers a reflective platform and a repository of practical strategies that can be adapted in similar contexts. For policymakers and educational institutions, it highlights the pressing need for structural reforms in teacher education, curriculum design, and resource allocation. For researchers, it identifies fertile ground for future inquiry, including longitudinal studies on the long-term effects of adaptive strategies, comparative analyses across regions, and explorations of technology-enhanced learning for deaf students.

This study thus aims to examine three interconnected domains: (1) the specific challenges English teachers face in delivering instruction to hearing-impaired students, (2) the adaptive strategies they develop and implement to overcome instructional and communication barriers, and (3) their perceptions of the effectiveness and limitations of these strategies. By drawing on qualitative interviews with experienced English teachers from several *Sekolah Luar Biasa* (SLBs) and inclusive schools in urban Indonesian contexts, the research seeks to produce a nuanced picture of inclusive English education as it is practiced on the ground.

Ultimately, this study contributes to ongoing efforts to make Indonesia's education system more inclusive, equitable, and responsive to the diverse needs of learners. By centering the experiences of teachers and their adaptive practices, it not only addresses a critical research gap but also provides actionable insights for building a more robust infrastructure for deaf-inclusive English education. In doing so, the research aspires to ensure that hearing-impaired students are no longer left on the margins of English learning but are instead empowered to participate fully in academic and social life, locally, nationally, and globally.

METHODS

This study adopted a qualitative research design to explore the lived experiences of teachers instructing deaf students in English at *Sekolah Luar Biasa* (SLB) and inclusive secondary

schools in Indonesia. Qualitative inquiry was chosen because it enables an in-depth understanding of participants' perspectives, practices, and the sociocultural contexts that shape them (Creswell & Plano, 2023). The design is particularly suitable for capturing the nuances of human interaction and meaning making, especially when examining complex issues such as inclusive education and deaf pedagogy, where standardized quantitative measures often fail to capture the richness of teacher agency, strategies, and constraints.

The study followed a phenomenological orientation to foreground teachers lived experiences in inclusive classrooms. Phenomenology allows researchers to capture not just observable practices but also the underlying beliefs, challenges, and motivations that influence teaching (Vagle, 2018). By focusing on teachers' narratives, this approach highlights how they navigate structural limitations, adapt instructional strategies, and negotiate between policy expectations and classroom realities. Semi-structured interviews were selected as the primary data collection method because they provide flexibility to probe individual experiences while maintaining comparability across participants (Al Balushi, 2018).

Participants consisted of English teachers from SLB schools and inclusive secondary schools across Yogyakarta and Central Java. A purposive sampling technique was employed to ensure that respondents had direct experience teaching English to deaf students, aligning with the study's objectives (Creswell & Plano, 2023). Purposive sampling allows the selection of information-rich cases that maximize depth of understanding. A total of ten teachers participated, representing a diversity of backgrounds in terms of teaching experience, institutional support, and prior exposure to special needs pedagogy. This variation was critical to capturing a wide spectrum of adaptive strategies and challenges, which enhances the credibility and transferability of the findings (Zairul, 2025). Table 1 below presents the participants' demographic such as; gender, years of experience, type of school.

TABLE 1 | List of participants

Participant (Pseudonym)	Gender	Years of Teaching Experience	School Type	Training in Special Needs/Deaf Pedagogy (Yes/No)	Notable Practices / Notes
Angela	F	12	SLB (Urban)	Yes (last training 2014, outdated)	Designs own materials, uses games, emphasizes peer support
Bella	F	8	SLB (Urban)	No	Uses color-coded markers (words, pronunciation, meaning), contextual learning
Charly	M	15	SLB (Urban)	No	Uses repetition, seating arrangements for visibility
Dennies	M	10	SLB (Urban)	No	Combines bilingual explanation and visuals
Ellen	F	6	SLB (Urban)	No	Focuses on visual aids, frequent review, vocabulary reinforcement
Farhat	F	7	SLB (Urban)	No	Employs multimodal strategies (sign + spoken + realia)

As presented in [Table 1](#), this study involved six teachers as the participants, each representing a different special education school in Indonesia. The diversity of schools was intentionally considered to ensure a wide range of perspectives and experiences related to teaching English for deaf students. This variation allowed the researcher to gain more comprehensive insights into the challenges and strategies used by teachers across different educational contexts.

Semi-structured interviews were the primary source of data. Each interview lasted between 45 and 60 minutes and was conducted in Bahasa Indonesia to ensure comfort and clarity for participants. An interview guide was developed, covering themes such as teaching experiences, instructional challenges, communication strategies, media use, curriculum adaptation, and professional support. However, participants were also encouraged to expand beyond the guiding questions, allowing new themes to emerge organically. Interviews were audio-recorded with participants' consent, transcribed verbatim, and anonymized to protect confidentiality, in accordance with ethical guidelines. To enhance validity, follow-up communications were conducted with several participants for clarification and member checking ([Creswell & Plano, 2023](#)).

Data analysis followed the six-phase thematic analysis framework proposed by ([Zairul, 2025](#)) to ensure transparency, consistency, and rigor. The process began with repeated reading of the transcripts to achieve familiarity, followed by generating initial codes. Coding was carried out manually and iteratively using ([Lungu, 2022](#)) methods, which distinguish between open coding, axial coding, and selective coding. This coding process enabled both descriptive categorization and interpretive synthesis, allowing patterns to be identified across cases ([Ahmed et al., 2025](#)). Axial coding further helped to link categories, such as connecting communication barriers to instructional adaptations, while selective coding led to the integration of these categories into broader themes, such as systemic constraints, teacher resilience, and policy gaps.

To ensure trustworthiness, multiple strategies were applied. Credibility was established through member checking, triangulation across different participants, and peer debriefing with fellow researchers. Transferability was addressed by providing thick descriptions of participants' contexts and teaching environments. Dependability and confirmability were supported by maintaining an audit trail of coding decisions and reflexive notes throughout the analysis ([Zairul, 2025](#)). Reflexivity was critical, as the researcher acknowledged their positionality as a university lecturer in English education, which provided both insider knowledge and potential biases that required careful management.

The research adhered to ethical principles of informed consent, voluntary participation, and confidentiality. All participants were provided with clear explanations of the study's aims, data usage, and their rights to withdraw at any stage. Ethical protocols which emphasize respect, transparency, and responsibility in educational research. Pseudonyms were used in transcripts and reporting to protect participants' identities.

Combining contemporary methodological frameworks with classical foundations strengthened the study's rigor and credibility. [Creswell & Plano \(2023\)](#), provided up-to-date tools for qualitative inquiry, coding, and case study validity. This hybrid approach ensured that the study met current academic standards while respecting the historical contributions of qualitative methodology.

RESULTS AND DISCUSSION

Challenges in Teaching English to Deaf Students in Indonesian SLB

One of the most prominent findings in this study is the significant foundational language deficit faced by deaf students in Indonesian *Sekolah Luar Biasa* (SLB), which severely affects their English as a Foreign Language (EFL) acquisition. Teachers consistently reported that students entered secondary school with underdeveloped proficiency in both Bahasa Indonesia and sign language.

Transcript 1 (Ellen reflected): *"Many of them were born deaf, they lack any language input from the beginning, so their vocabulary is very limited."*

The transcript shows that Ellen stated how her students' limited vocabulary stems from their early language deprivation. Because they were born deaf and did not receive sufficient linguistic input from birth, their overall language foundation, both in sign and spoken form, remains underdeveloped, which directly hinders English vocabulary learning.

Transcript 2 (Farhat reflected): *"My students have limited vocabulary in Indonesian, let alone abstract words."* Farhat emphasized that students' difficulties in understanding abstract English words are rooted in their weak command of Indonesian. Without a strong first language base, transferring meaning to English becomes a major challenge.

The inconsistency between English spelling and pronunciation posed additional barriers for students who relied on visual and speechreading strategies. Transcript 3 (Bella noted): *"In English, the written form and pronunciation are different, right? Meanwhile, they rely on lip reading."* This transcript indicates Bella's awareness of how the inconsistency between English spelling and pronunciation confuses deaf students, who depend heavily on visual cues and lip reading to comprehend lessons.

Transcript 4 (Dennies reflected): *"English spelling and pronunciation often don't match, so students get confused."* Dennies confirmed similar challenges to Bella's experience, highlighting how orthographic irregularity in English makes it difficult for students to connect written and spoken forms. Teachers addressed these challenges by simplifying input and providing visual aids. Transcript 5 (Charly explained): *"Some students still have hearing ability, so I only need to repeat a few times. But for totally deaf students, it's much harder, even the next day, they still might not be able to pronounce it."* Charly's statement shows how differing levels of residual hearing among students require varied teaching repetition. Students with complete hearing loss need continuous repetition and visual reinforcement to retain knowledge.

Transcript 1 (Ellen added): *“Sometimes they don’t understand the writing, but when they see the picture, they understand instantly.”* This transcript illustrates Ellen’s realization that visual aids play a crucial role in comprehension for deaf learners, who respond better to images than text alone.

Teachers consistently identified weak retention as a challenge. Transcript 6 (Angela shared): *“Sometimes they can understand it in the classroom context, but they forget it in real life.”* While in transcript 5 (Charly remarked): *“What we teach today might be forgotten by the next day.”* Transcript 1 Ellen highlighted the need for repetition: *“This process requires repetition. Hopefully, they will remember.”* Also transcript 2 Farhat linked memory to engagement: *“I try to make them active, both cognitively and physically, so that learning becomes meaningful.”*

In response to these difficulties, teachers employed multimodal approaches to overcome communication breakdowns. Transcript 2 (Farhat recalled): *“If I only use spoken words, they get confused. I once tried giving instructions verbally only, then covered my own ears, and realized how hard it was.”* Transcript 3 Bella stressed visual accessibility: *“Lip reading only helps when they can clearly see how I move my mouth. I need to make sure they can see me all the time.”* Transcript 6 Angela described her varied toolkit: *“I use Canva worksheets, color markers, puzzles, everything combined, because one method doesn’t work.”*

Many teachers expressed feelings underprepared due to limited or outdated training. Angela explained: *“Actually, I’m a specialized teacher myself. I just have the extra duty of teaching English. So, we try to understand the content ourselves.”* She added: *“My last training was in 2014. At the time, the content was relevant. But now, it’s outdated.”* Farhat admitted: *“I don’t know if my methods are pedagogically correct... I rely solely on personal experience.”* Bella shared an improvisation: *“I use a black marker for the English words, a blue marker for the pronunciation, and a red marker for the meaning.”*

Charly identified language itself as the biggest obstacle: *“The biggest challenge is language. Even in Indonesian, they don’t fully understand.”* Dennies reinforced this: *“Since they can’t hear, they can’t learn language the way most children do, through sound.”* Teachers often resorted to bilingual explanations and visual scaffolds, but gaps persisted when neither Indonesian nor BISINDO was fully developed.

Teachers reported weak school–family communication and limited external collaboration. Angela stated: *“I always ask them to notify parents.”* Bella admitted: *“I haven’t had direct communication with parents.”* Farhat acknowledged: *“Support is limited. Some parents don’t know English or don’t have time.”* Angela also noted the absence of professional networks: *“We validate ourselves by consulting with other English teachers.”*

Teachers had to develop their own resources due to the lack of tailored textbooks. Bella stated: *“There are no English textbooks at all... the teacher really has to create the materials.”* Angela called for change: *“Please consider making teaching materials for teachers, tailored for deaf, blind, or physically disabled students.”*

The findings reveal a multi-layered set of barriers to EFL learning for deaf students in Indonesian SLBs, rooted in early

language deprivation, phonological-orthographic mismatch, cognitive retention issues, and systemic gaps in teacher preparation and resources.

Consistent with [Hartman et al., \(2019\)](#), the teachers’ accounts confirm that delayed or absent early L1 development, whether in spoken or signed form, creates long-term challenges for additional language learning. Without a fully formed linguistic base, students lack the conceptual scaffolding necessary for vocabulary acquisition and grammar comprehension in EFL. This aligns with [Hamilton et al., \(2024\)](#), who stresses that incomplete early language access reduces readiness for L2 or L3 learning.

The mismatch between English spelling and pronunciation, as described by [Almusawi, \(2019\)](#), assertion that non-transparent grapheme–phoneme correspondence is particularly problematic for deaf learners. Because many rely on visual input and lip reading, irregular orthography complicates decoding and retention. Teachers’ adaptations, such as visual aids, color coding, and repeated exposure, are consistent with multimodal literacy strategies recommended by ([Prystiananta & Noviyanti, 2025](#)).

The difficulty in retaining vocabulary and concepts, reported by [Nugroho & Lintangari, \(2022\)](#), is well-documented in deaf education literature. [Marschark & Hauser \(2006\)](#), note that deaf learners often require more frequent and varied reinforcement to consolidate learning, especially when L1 is weak. Farhat’s emphasis on active engagement reflects constructivist approaches, where physical and cognitive participation can improve long-term recall.

Teachers’ use of Canva worksheets, realia, and physical movement aligns with the Total Communication (TC) philosophy, which advocates integrating visual, tactile, and auditory inputs to maximize comprehension ([Sugiarni et al., 2024](#)). However, the reliance on personal improvisation, rather than structured methodology, risks inconsistency in quality across classrooms, a concern echoed by ([Patzak & Zhang, 2025](#)).

The teachers’ accounts reveal a systemic shortage of specialized EFL training for educators of deaf students, echoing [Ahmad & Khasawneh \(2021\)](#), findings on the mismatch between teacher assignments and formal preparation. Without ongoing training, teachers depend on trial-and-error methods, leading to varied outcomes. The outdated nature of existing training programs, as Angela described, underscores the need for continuous professional learning that integrates current research on deaf education and language acquisition.

The minimal parental involvement and lack of institutional partnerships reported in this study mirror patterns found by [Motshusi et al., \(2024\)](#), in other inclusive settings. Without active home–school–community engagement, students’ learning remains confined to classroom boundaries, limiting reinforcement and contextualization. [Ostovar-Nameghi and Sheikahmadi \(2016\)](#), emphasize that professional isolation among teachers reduces innovation and the sharing of best practices, an issue visible here.

The complete absence of deaf-friendly EFL textbooks forces teachers into material self-production, which, while creative, is unsustainable in the long term. [Onyishi and Sefotho \(2020\)](#), argue that institutional provision of

differentiated materials is a prerequisite for true inclusion; without it, inclusive education remains aspirational rather than practical.

Strategies Used by Teachers in Teaching the Deaf

English teachers in Indonesian *Sekolah Luar Biasa* (SLB) adopt highly visual, multimodal, and interactive strategies to compensate for the auditory barriers faced by deaf students in learning English as a Foreign Language (EFL). Farhat explained: *"I use flashcards to introduce vocabulary... when discussing fruits, I show the actual fruit, like apples or bananas."* Angela emphasized color and visual aids: *"I usually use colored markers... I use full visual support, text and images, to help them understand"*. Bella reinforced contextual learning: *"I explain its color, taste, green on the outside, orange inside, whether sweet or sour."* These practices illustrate the centrality of realia, visuals, and descriptive detail in facilitating comprehension.

Teachers combined spoken language, sign language, and demonstrations to clarify meaning. Farhat noted: *"If I only use spoken words, they get confused... I always give instructions clearly, verbally, with sign language, and demonstrations."* Bella highlighted classroom visibility *"If they can't see me, they won't know what's going on."* These strategies show deliberate classroom arrangements and redundant communication to ensure understanding.

Despite the effectiveness of these approaches, no teacher interviewed had received specialized EFL training for deaf learners. Farhat admitted: *"I've never received any specific training... I rely solely on personal experience."* Angela added: *"We validate ourselves by consulting with fellow English teachers."*

Lessons were adapted to students' everyday experiences and limited vocabulary exposure. Angela said: *"I focus on teaching functional English rather than strictly following textbook content."* Farhat explained: *"What I teach is still limited to basic vocabulary... daily life."* Angela also described structural scaffolding: *"I use simplified structures like 'Subject + to be + noun' with visual support."* Bella stressed repetition: *"We repeat things often... repetition helps it stick."*

Peer support was encouraged to reinforce understanding. Angela shared: *"If student A finishes first and gets it right, I ask them to help others."* Bella said: *"I ask that student to explain it to their friend. So, it becomes a kind of peer review."* Visibility was prioritized: *"We place them in the front row... visibility helps them stay engaged."*

Teachers reported creating all instructional materials themselves. Bella stated: *"There are no English textbooks at all... so the teacher really has to create the materials themselves."* Angela noted: *"The bright ones have no problem with videos, but the others only understand the visuals, not the content."*

Games and movement-based activities supplemented instruction. Angela said: *"I design custom boards where students use dice and circle vocabulary in turns."* Charly added: *"Even in Indonesian, they don't fully understand... so I tried using a game. Sometimes what we teach is forgotten by the next day."*

Angela differentiated materials for individual learners: *"Sometimes I differentiate for just one or two students in a class of five."* Dennies focused on functional daily content: *"I taught them how to make tea."* Teachers often switched between sign language, Bahasa Indonesia, and local dialects to bridge gaps.

Assessment methods avoided heavy reliance on auditory input. Angela used: *"Cloze paragraphs with pictures"* and *"fill-in-the-blank exercises."* Farhat preferred: *"Matching and labeling" using tactile tools*. Emotional check-ins were done with smiley-faced icons or observation.

Teachers navigated diverse proficiency levels in the same classroom. Charly explained: *"One student is still learning 'What's your name?' and another is already doing prepositions."* Ellen added: *"I don't strictly follow the curriculum because it's too general."*

Effectiveness of Strategies in Promoting Language Acquisition

The acquisition of English among deaf students in Indonesia's *Sekolah Luar Biasa* (SLB) shows a pattern of gradual progress, with stronger development in concrete and functional vocabulary rather than abstract grammar or complex linguistic structures. Teachers consistently adopt a communicative, context-driven approach, prioritizing practical relevance over strict adherence to textbooks. Angela remarked: *"I focus on teaching functional English rather than strictly following textbook content."* Farhat explained: *"What I teach is still limited to basic vocabulary, especially words related to their everyday lives."*

To reinforce vocabulary, teachers rely heavily on visual and tactile scaffolding aligned with the Total Communication (TC) philosophy. Angela described: *"I use colored markers and word search worksheets."* Farhat added: *"Flashcards, realia, and demonstrations"* were core to her lessons. However, learner diversity within the same classroom creates challenges. Charly stated: *"Even within the same 10th-grade class, one student is still learning 'What's your name?' and another is already doing prepositions."* Dennies explained his step-by-step approach: *"First translation, then illustrations, and gradually introduce vocabulary."* Farhat emphasized adaptability: *"We must truly find the right strategy for each student's condition."* Angela also noted: *"Even though they're all deaf, their abilities vary greatly... Some are quick to grasp material, while others struggle with understanding English."*

Peer assistance plays a significant role. Angela shared: *"If student A finishes first and gets it right, I ask them to help others."* Bella confirmed: *"I ask that student to explain it to their friend."* Angela reflected on its impact: *"Students who help others feel proud... They also become more engaged in class."* Yet, systemic gaps remain. Angela admitted: *"Our training was long ago and outdated... But peer help works instantly."* Farhat added: *"I've never received any specific training for teaching English to deaf students."*

For assessment, teachers use visually oriented, performance-based strategies. Angela reported: *"Students enjoy word search activities and cloze paragraph tasks."* Farhat said: *"Hands-on activities like flashcard matching or sticking pictures help them retain information better when they can interact directly."*

Game-based learning also features prominently. Angela described: *"We use dice, and they take turns circling vocabulary on the board"* (Snakes and Ladders). Charly shared: *"I draw on the floor to teach prepositions... they have to jump in those directions."* Despite these innovations, teachers note misalignment with national standards. Angela said: *"We're still experimenting, what worked last year might not work this year."* Dennies remarked: *"There are no textbooks or standardized tools for English in SLB."*

Language transfer is another difficulty. Angela remarked: *"Even when students enjoy the activities, they rarely transfer that vocabulary into spontaneous use."* Dennies confirmed: *"They can match a word with a picture, but they don't use it outside the activity."* Ellen explained: *"They lack any language input from the beginning."* Farhat added: *"They don't know the Indonesian word, so I have to teach both it and the English translation from scratch."*

To address these gaps, Farhat uses repetition: *"Before starting new material, I always review, ask, 'What did we learn yesterday?'"* Finally, resource shortages hinder instruction. Angela stated: *"Technology has no limits, but it's hard to find what fits the students' needs."* Bella was blunter: *"There are no English textbooks at all in SLB."*

Challenges in Teaching English to Deaf Students in Indonesian SLB

The findings reveal that EFL instruction for deaf students in Indonesian SLBs is built on visual, multimodal, and functional strategies, hallmarks of the Total Communication (TC) approach. This mirrors [Marschark et al., \(2015\)](#), assertion that combining spoken, signed, written, and visual modes improves comprehension for deaf learners. Teachers' use of realia, color coding, descriptive explanations, and gesture that multimodal strategies support the acquisition of abstract language among learners with limited L1 proficiency.

A significant concern is the absence of specialized training for teaching EFL to deaf students. This forces teachers to rely on experience and peer advice, echoing [Sari et al., \(2022\)](#), findings about the gap between inclusive education policy and classroom practice. Without targeted professional development, instructional quality depends heavily on individual creativity rather than evidence-based practice.

The emphasis on functional vocabulary and real-life contexts reflects [Yildiz, Y., & Celik, B. \(2020\)](#), recommendations for contextual learning to overcome linguistic delays. Simplified structures, repetition, and scaffolding help compensate for gaps in early language exposure, while contextual teaching improves retention and relevance.

Peer-assisted strategies reported here resonate with [Sahara et al., \(2024\)](#), findings that structured peer interaction benefit both cognitive development and social integration in deaf learners. Classroom seating arrangements that maximize visibility further align with TC principles.

The lack of EFL materials for deaf learners is a critical gap. [Weiss et al., \(2021\)](#), stressed that accessible curricula and differentiated materials are essential for inclusive success. The teachers' need to design all their own materials highlights

systemic neglect in resource provision, which can contribute to burnout.

Interactive games and movement-based learning, as described by [Abdoulqadir & Loizides \(2025\)](#), call for multisensory approaches and individualized learning plans. However, the continued problem of short-term retention indicates the need for structured review cycles and spaced repetition strategies.

Differentiation for individual learners reflects [Ituma \(2025\)](#), advocacy for responsive teaching. Switching between sign language, Bahasa Indonesia, and local dialects reflects [Rasman \(2018\)](#), findings on the importance of multilingual approaches in inclusive classrooms.

Performance-based and visual assessment methods, such as matching and picture-supported cloze tests, correspond to [Luft, \(2020\)](#), recommendation for non-auditory evaluation tools. These methods respect deaf learners' preferences and allow them to demonstrate knowledge without oral output.

The vast differences in student proficiency, noted by [Patel & Kim \(2024\)](#), align with Vygotsky's Zone of Proximal Development (ZPD) theory, which calls for instruction tailored to each learner's readiness level. Without flexible curricula, these disparities can limit collective progress in mixed-ability classrooms.

Strategies Used by Teachers in Teaching the Deaf

The findings from the interviews with English teachers in Indonesian *Sekolah Luar Biasa* (SLB) highlight a clear emphasis on highly visual, multimodal, and interactive instructional strategies to support deaf students in learning English as a Foreign Language (EFL). Teachers reported employing flashcards, real objects (realia), images, color-coded markers, and descriptive explanations to enhance comprehension, consistent with the literature emphasizing the importance of visual reinforcement in deaf education ([Marschark & Hauser, 2006](#)). For example, Farhat's practice of showing real fruits and Bella's detailed description of color, taste, and texture reflect the adoption of contextualized, multimodal strategies to facilitate vocabulary acquisition, a finding aligned with research advocating for sensory-rich, experiential learning for hearing-impaired students ([Baliber-Duallo, 2025](#)).

The integration of spoken language, sign language, and demonstrations further illustrates the use of Total Communication (TC) approaches in practice, supporting findings in prior studies that multimodal communication improves comprehension and reduces frustration among deaf learners ([Shaver et al., 2014](#)). Teachers' attention to classroom visibility and deliberate seating arrangements emphasizes the importance of environmental and instructional scaffolding in EFL classrooms for hearing-impaired students, echoing [Guardino and Fullerton \(2010\)](#) findings that visual accessibility and classroom organization significantly influence engagement and understanding.

Despite these adaptive practices, the interviews revealed a notable gap: none of the teachers had received specialized EFL training for deaf learners. Teachers relied primarily on personal experience and peer consultation, which underscores the persistent challenge identified in the literature regarding the insufficient professional development opportunities for

educators of hearing-impaired students in Indonesia (Villarente, 2024). This highlights a critical need for structured training programs to enhance teachers' pedagogical knowledge and confidence, enabling them to implement evidence-based strategies more systematically.

Lesson adaptation emerged as another central theme. Teachers emphasized functional English, simplified structures, repetition, and scaffolding to match students' everyday experiences and limited vocabulary exposure. These approaches correspond with the literature on differentiated instruction and individualized learning plans (ILPs), which suggest tailoring content and instructional strategies to students' abilities and needs improves comprehension and engagement (Hossain, 2024). Similarly, the reported use of peer-assisted strategies and front-row seating supports prior findings that cooperative learning and peer support enhance comprehension, engagement, and social interaction for hearing-impaired learners (Guardino & Fullerton, 2010).

Another important finding is the teachers' reliance on self-created instructional materials and interactive, game-based learning activities. While these approaches foster engagement and contextual learning, retention remains challenging, as teachers reported frequent forgetting of previously taught content. This aligns with research emphasizing that multimodal strategies, when combined with repetition and scaffolded instruction, are more likely to support long-term retention in deaf learners (Hauser & Marschark, 2008). The use of differentiated materials and frequent code-switching between sign language, Bahasa Indonesia, and local dialects demonstrates an adaptive, student-centered approach responsive to individual needs, a practice consistent with inclusive education principles (Yow et al., 2018).

Assessment practices also reflected a shift from auditory-dependent methods toward visually and tactilely oriented tools, including cloze paragraphs, matching exercises, and icon-based emotional check-ins. These strategies resonate with recommendations from (Woolley, 2011), regarding formative, multimodal assessment and feedback, which support individualized monitoring of learning progress and foster student self-awareness.

Finally, teachers' navigation of diverse proficiency levels within a single classroom highlights the complexity of implementing inclusive EFL instruction in SLB contexts. The variability in students' prior knowledge and skill levels necessitates flexible lesson planning and real-time differentiation, reinforcing findings that effective inclusive education requires ongoing adaptation, collaboration, and sensitivity to learners' unique profiles (Guardino & Fullerton, 2010).

Effectiveness of Strategies in Promoting Language Acquisition

The finding reveals that teachers' emphasis on functional vocabulary over abstract grammar aligns with Kaharuddin (2018), recommendation to focus on practical language use in special education contexts. Such an approach helps bridge the gap caused by many students delayed or fragmented L1 acquisition, a challenge well-documented in deaf education.

In Indonesia, although inclusive education policies exist, implementation in SLB settings often lacks structured linguistic support and tailored resources (Ediyanto et al., 2021).

The heavy reliance on visual, tactile, and multimodal strategies reflects principles of the Total Communication approach, which combines sign, visual aids, and kinesthetic learning to enhance understanding for deaf students (Wainscott & Spurgin, 2024). However, the wide range of abilities in a single class requires constant differentiation, an issue noted as central to the success of Individualized Learning Plans (ILPs) (Ardenlid, 2025).

The variation in students' linguistic backgrounds, ranging from Bahasa Indonesia to BISINDO, home signs, or local dialects, means many lack a stable L1 foundation. This significantly complicates L2 English learning and, in some cases, effectively turns English into a third or fourth language. Without solid L1 development, additional languages face substantial acquisition barriers (Septiani et al., 2024).

Peer-assisted learning emerged as a low-cost yet impactful scaffold. Vygotsky's Zone of Proximal Development (ZPD) framework, referenced by Lasmawan & Budiarta (2020), supports this practice, showing that learners often progress through interaction with more capable peers. Moreover, deaf learners often co-construct communication norms, making peer mediation particularly effective in multilingual or multi-sign classrooms (Pizzo, 2016). Nevertheless, the teachers' reliance on such strategies reveals systemic shortcomings in teacher preparation, insufficient training investment for special needs educators in Indonesia (Al Aqsha & Emilzoli, 2024).

In assessment, the teachers' preference for visual-spatial and kinesthetic activities is well-grounded in the sensory strengths of deaf learners such as integrating movement and tactile engagement into evaluation (Alenizi, 2019). However, the absence of standardized SLB-specific English assessments leaves these innovations without formal recognition or scalability.

Language transfer challenges, noted by multiple teachers, suggest a gap between activity-based recognition and real-world application. This reflects without deep semantic processing, vocabulary knowledge remains shallow and context bound (Howerton-Fox & Falk, 2019).

Finally, the discussion on resources and sustainability points to a larger systemic issue. Teachers' ingenuity in creating materials is admirable, but inclusion cannot rely on individual improvisation, it must be supported by structured training, curriculum development, and institutional collaboration (Woodcock et al., 2022). Without such systemic backing, the burden remains on under-supported educators, jeopardizing the long-term success of inclusive EFL instruction in SLBs.

CONCLUSION

This study provides three key contributions to the field of deaf-inclusive English education in Indonesia. First, it offers empirical insight into teachers' lived experiences, revealing the challenges of linguistic delays, communication barriers,

limited training, scarce resources, and rigid curricula that hinder effective EFL instruction in *Sekolah Luar Biasa* (SLB) and inclusive secondary schools. Second, it presents a catalog of adaptive strategies, including flashcards, realia, visual storytelling, kinesthetic games, peer-assisted learning, and optimized classroom arrangements, highlighting teachers' resilience and creativity in fostering engagement and vocabulary acquisition despite systemic constraints. Third, it delivers policy-level recommendations for sustainable inclusive education, advocating for curricula grounded in functional vocabulary, teacher training in deaf pedagogy and bimodal bilingual strategies, visual and performance-based assessments, accessible infrastructure, and accountable policy implementation.

Based on the findings of this study, it is recommended that deaf-inclusive English education in Indonesia adopt a holistic approach addressing teachers, students, institutions, and policymakers. Teachers should engage in professional development focused on deaf pedagogy, Total Communication, and bimodal bilingual strategies while applying adaptive techniques such as flashcards, realia, visual storytelling, kinesthetic games, and peer-assisted learning to enhance engagement and comprehension. Students should benefit from highly visual, interactive, and contextualized activities, reinforced through repetition, scaffolding, and cooperative learning to support vocabulary acquisition, functional English skills, and confidence. Inclusive education institutions are urged to provide accessible infrastructure, including captioned media, visual teaching aids, adaptive technology, and classroom designs optimized for visibility, while ensuring systematic support, resource allocation, and curricula emphasizing functional vocabulary and performance-based assessment. Policymakers must translate inclusive education policies into actionable frameworks with clear funding, monitoring, and accountability mechanisms, fostering collaboration among ministries, schools, universities, and advocacy groups. Finally, further research is encouraged to explore the long-term effectiveness of adaptive strategies, technology integration, and additional support mechanisms, building a robust evidence base to strengthen inclusive EFL instruction. Collectively, these measures aim to create sustainable, equitable, and high-quality English learning opportunities for hearing-impaired students in Indonesia.

ACKNOWLEDGEMENTS

This research was supported by a research grant from Universitas Muhammadiyah Yogyakarta. The authors would also like to acknowledge the assistance of ChatGPT, which was used to improve the clarity and refinement of the language in this manuscript.

REFERENCES

Abdoulqadir, C., & Loizides, F. (2025). Interaction, artificial intelligence, and motivation in children's speech learning and rehabilitation through digital games: A

systematic literature review. *Information (Switzerland)*, 16(7), 599.

<https://doi.org/10.3390/INFO16070599/S1>

Adeduyigbe, A. M., Adeduyigbe, A. E., & Tijani, B. E. (2024). Addressing students with hearing impairment's current state and future needs: Reforming an inclusive science education. *International Journal of Studies in Inclusive Education*, 1(2), 12–15.

<https://doi.org/10.38140/IJSIE.V1I2.1448>

Ahmad, M., & Khasawneh, S. (2021). Problems teaching English to deaf students. *Indonesian Journal of Creative Counseling*, 1(2), 32–42.

<https://doi.org/10.47679/ijcc.v1i2.107>

Ahmed, S. K., Mohammed, R. A., Nashwan, A. J., Ibrahim, R. H., Abdalla, A. Q., M. Ameen, B. M., & Khdhir, R. M. (2025). Using thematic analysis in qualitative research. *Journal of Medicine, Surgery, and Public Health*, 6, 100198.

<https://doi.org/10.1016/J.GLMEDI.2025.100198>

Al Aqsha, I., & Emilzoli, M. (2024). Competence of teachers to support special needs children in our dream Indonesia. *Eduvest - Journal of Universal Studies*, 4(5), 3980–3992.

<https://doi.org/10.59188/EDUVEST.V4I5.1173>

Al Balushi, K. (2018). The use of online semi-structured interviews in interpretive research. *International Journal of Science and Research (IJSR)*, 7(4), 726–732. <https://doi.org/10.21275/ART20181393>

Alanazi, M. (2021). Communicating with deaf students in inclusive schools: insights from Saudi university faculty. *Eurasian Journal of Educational Research*, 95(95), 188–209.

<https://doi.org/10.14689/EJER.2021.95.11>

Alenizi, M. A. K. (2019). Effectiveness of a program based on a multi-sensory strategy in developing visual perception of primary school learners with learning disabilities: A contextual study of Arabic learners. *International Journal of Educational Psychology*, 8(1), 72–104.

<https://doi.org/10.17583/IJEP.2019.3346>

Almusawi, H. M. (2019). Determinants of spelling proficiency in hearing and deaf graduate students: The presentation of medial glottal stop. *Ampersand*, 6, 100050.

<https://doi.org/10.1016/J.AMPER.2019.100050>

Ardenlid, F. (2025). Differentiated instruction for gifted students and their peers in Swedish mixed-ability classrooms: teachers' principles and practices. *Cogent Education*, 12(1).

<https://doi.org/10.1080/2331186X.2025.2520560:WGROUP:STRING:PUBLICATON>

Baliber-Duallo, R. (2025). Vocabulary acquisition in deaf education: pedagogical approaches for supporting learners with hearing impairments. *Cognizance Journal of Multidisciplinary Studies*, 5(4), 1143–1150. <https://doi.org/10.47760/COGNIZANCE.2025.V05I04.040>

Bintoro, T., Kusmawati, A. P., & Dewi, R. S. (2023). The teacher strategies in teaching sign language for deaf

- students in special schools Jakarta. *Cogent Education*, 10(2).
<https://doi.org/10.1080/2331186X.2023.2258294>
- Bowen, S. K., & Probst, K. M. (2023). Deaf and hard of hearing students with disabilities: An evolving landscape. *Education Sciences*, 13(7).
<https://doi.org/10.3390/EDUCSCI13070752>
- Creswell, J. W., & Plano Clark, V. L. (2023). Revisiting mixed methods research designs twenty years later. *Handbook of mixed methods research designs*, 1(1), 21–36.
<https://www.torrossa.com/en/resources/an/5730651#page=56>
- Ediyanto, E., Setiawan, A., Handaka, I. B., Rofiah, N. H., & Suhendri, S. (2021). Implementation of inclusive education in learning process at senior high school Malang city, Indonesia. *Indonesian Journal of Disability Studies*, 8(1), 179–189.
<https://doi.org/10.21776/UB.IJDS.2021.008.01.14>
- Guardino, C. A., & Fullerton, E. (2010). Changing behaviors by changing the classroom environment. *TEACHING Exceptional Children*, 42(6), 8–13.
<https://doi.org/10.1177/004005991004200601>
- Hamilton, C., Schulz, J., Chalmers, H., & Murphy, V. A. (2024). Investigating the substantive linguistic effects of using songs for teaching second or foreign languages to preschool, primary and secondary school learners: A systematic review of intervention research. *System*, 124, 103350.
<https://doi.org/10.1016/J.SYSTEM.2024.103350>
- Hartman, M. C., Nicolarakis, O. D., & Wang, Y. (2019). Language and literacy: Issues and considerations. *Education Sciences*, 9(3).
<https://doi.org/10.3390/EDUCSCI9030180>
- Hauser, P. C., & Marschark, M. (2008). What we know and what we don't know about cognition and deaf learners. *Deaf Cognition: Foundations and Outcomes*.
<https://doi.org/10.1093/ACPROF:OSO/9780195368673.003.0016>
- Hiver, P., Al-Hoorie, A. H., Vitta, J. P., & Wu, J. (2024). Engagement in language learning: A systematic review of 20 years of research methods and definitions. *Language Teaching Research*, 28(1), 201–230.
<https://doi.org/10.1177/13621688211001289>
- Hossain, K. I. (2024). Literature-based language learning: Challenges, and opportunities for English learners. *Ampersand*, 13, 100201.
<https://doi.org/10.1016/J.AMPER.2024.100201>
- Howerton-Fox, A., & Falk, J. L. (2019). Deaf Children as 'English learners': The psycholinguistic turn in deaf education. *Education Sciences* 2019, 9(2), 133.
<https://doi.org/10.3390/EDUCSCI9020133>
- Ituma, M. G. (2025). The use of differentiated instruction to achieve culturally responsive teaching. *Open Journal of Educational Research*, 5(1), 13–30.
<https://doi.org/10.31586/OJER.2025.1234>
- John W. Creswell, & J. David Creswell. (2023). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Sage Publications, Inc.
- Kaharuddin, A. (2018). The communicative grammar translation method: a practical method to teach communication skills of English. *Eternal (English, Teaching, Learning, and Research Journal)*, 4(2), 232.
<https://doi.org/10.24252/ETERNAL.V42.2018.A8>
- Lasmawan, I. W., & Budiarta, I. W. (2020). Vygotsky's zone of proximal development and the students' progress in learning (a heutagogical bibliographical review). *JPI (Jurnal Pendidikan Indonesia)*, 9(4), 545.
<https://doi.org/10.23887/JPI-UNDIKSHA.V9I4.29915>
- Luft, P. (2020). Strengths-based reading assessment for deaf and hard-of-hearing students. *Psychology in the Schools*, 57(3), 375–393.
<https://doi.org/10.1002/PITS.22277>
- Lungu, M. (2022). The coding manual for qualitative researchers. *American Journal of Qualitative Research*, 6(1), 232–237.
<https://doi.org/10.29333/AJQR/12085>
- Marschark, M., & Hauser, P. C. (2006). *How Deaf Children Learn*. Oxford University Press, 1–6.
<https://global.oup.com/academic/product/how-deaf-children-learn-9780195389753>
- Marschark, M., Spencer, L. J., Durkin, A., Borgna, G., Convertino, C., Machmer, E., Kronenberger, W. G., & Trani, A. (2015). Understanding language, hearing status, and visual-spatial skills. *The Journal of Deaf Studies and Deaf Education*, 20(4), 310.
<https://doi.org/10.1093/DEAFED/ENV025>
- Motshusi, M. C., Ngobeni, E. T., & Sepeng, P. (2024). Lack of parental involvement in the education of their children in the foundation phase: Case of selected schools in the thabazimbi circuit. *Research in Educational Policy and Management*, 6(2), 21–41.
<https://doi.org/10.46303/REPAM.2024.20>
- Muller, M. (2022). Significance of hybrid learning model during covid-19 pandemic at higher education institution. *European Journal of Teaching and Education*, 4(2), 70–80.
<https://doi.org/10.33422/EJTE.V4I2.818>
- Mulyadi, A. W. E. (2017). Policy of inclusive education for education for all in Indonesia. *Policy & Governance Review*, 1(3), 201.
<https://doi.org/10.30589/PGR.V1I3.57>
- Njonge, T. (2023). Influence of psychological well-being and school factors on delinquency, during the covid-19 period among secondary school students in selected schools in Nakuru county: Kenya. *International Journal of Research and Innovation in Social Science (IJRISS)*, 1175–1189. <https://doi.org/10.47772/IJRISS>
- Nugroho, F. A., & Lintang Sari, A. P. (2022). Deaf students' challenges in learning English. *IJDS Indonesian Journal of Disability Studies*, 9(02), 217–224.
<https://doi.org/10.21776/UB.IJDS.2022.009.02.06>
- Onyishi, C. N., & Sefotho, M. M. (2020). Teachers' perspectives on the use of differentiated instruction in inclusive classrooms: Implication for teacher education. *International Journal of Higher Education*, 9(6), 136–150. <https://doi.org/10.5430/ijhe.v9n6p136>

- Ostovar-Nameghi, S. A., & Sheikahmadi, M. (2016). From teacher isolation to teacher collaboration: Theoretical perspectives and empirical findings. *English Language Teaching*, 9(5), 197. <https://doi.org/10.5539/elt.v9n5p197>
- Patel, D., & Kim, N. (2024). Differentiated instruction in mixed-ability classrooms: Experiences of special education teachers. *The Psychological Research in Individuals with Exceptional Needs*, 2(3), 13–20. <https://doi.org/10.61838/KMAN.PRIEN.2.3.3>
- Patzak, A., & Zhang, X. (2025). Blending teacher autonomy support and provision of structure in the classroom for optimal motivation: A systematic review and meta-analysis. *Educational Psychology Review*, 37(1). <https://doi.org/10.1007/S10648-025-09994-2>
- Pizzo, L. (2016). D/Deaf and hard of hearing multilingual learners: The development of communication and language. *American Annals of the Deaf*, 161(1), 17–32. <https://doi.org/10.1353/AAD.2016.0017>
- Prystiananta, N. C., & Noviyanti, A. I. (2025). Enhancing EFL instruction in special needs education: Integrating multimodal digital tools and deep learning strategies. *Voices of English Language Education Society*, 9(1), 96–108. <https://doi.org/10.29408/VELES.V9I1.29411>
- Rasman. (2018). To translanguage or not to translanguage? The multilingual practice in an Indonesian EFL classroom. *Indonesian Journal of Applied Linguistics*, 7(3), 687–694. <https://doi.org/10.17509/IJAL.V7I3.9819>
- Sahara, B., Mumpuniarti, M., Suwarjo, S., & Syamsuri, M. (2024). Perspectives of peers as a microsystem for supporting deaf students in inclusive elementary schools. *International Journal of Learning, Teaching and Educational Research*, 23(4), 298–319. <https://doi.org/10.26803/IJLTER.23.4.16>
- Salvaña, L., & Protacio, A. (2025). Navigating inclusive classrooms: English language teachers' narratives of teaching students with special needs. *Psychology and Education: A Multidisciplinary Journal*, 40(10), 1266–1282. <https://doi.org/10.70838/PEMJ.401002>
- Sari, Z. P., Sarofah, R., & Fadli, Y. (2022). The implementation of inclusive education in indonesia: challenges and achievements. *Jurnal Public Policy*, 8(4), 264. <https://doi.org/10.35308/JPP.V8I4.5420>
- Septiani, D., Zainal Rafli, Fathiatty Murtadho, & Tryana. (2024). The acquisition of word classes in the Indonesian language by students with hearing impairments. *Jurnal Sastra Indonesia*, 13(3), 252–263. <https://doi.org/10.15294/H6F1AS93>
- Shaver, D. M., Marschark, M., Newman, L., & Marder, C. (2014). Who is where? Characteristics of deaf and hard-of-hearing students in regular and special schools. *Journal of Deaf Studies and Deaf Education*, 19(2), 204–219. <https://doi.org/10.1093/deafed/ent056>
- Sugiarni, Widiastuti, D. E., & Tahrun. (2024). The implementation of Canva as a digital learning tool in English learning at vocational school. *English Learning Innovation*, 5(2), 264–276. <https://doi.org/10.22219/ENGLIE.V5I2.34839>
- Vagle, M. D. (2018). Crafting phenomenological research, second edition. *Crafting Phenomenological Research, Second Edition*, 1–198. <https://doi.org/10.4324/9781315173474>
- Villarente, S. V. D. (2024). Navigating challenges and implementing strategies: Teachers teaching deaf students in a higher education. <https://doi.org/10.22271/27103862.2024.v4.i1b.74>
- Wainscott, S. D., & Spurgin, K. (2024). Differentiating language for students who are deaf or hard of hearing: A practice-informed framework for auditory and visual supports. *Language, Speech, and Hearing Services in Schools*, 55(2), 473–494. https://doi.org/10.1044/2023_LSHSS-22-00088
- Weiss, S., Muckenthaler, M., Heimlich, U., Kuechler, A., & Kiel, E. (2021). Teaching in inclusive schools. Do the demands of inclusive schools cause stress? *International Journal of Inclusive Education*, 25(5), 588–604. <https://doi.org/10.1080/13603116.2018.1563834>
- Woodcock, S., Sharma, U., Subban, P., & Hitches, E. (2022). Teacher self-efficacy and inclusive education practices: Rethinking teachers' engagement with inclusive practices. *Teaching and Teacher Education*, 117, 103802. <https://doi.org/10.1016/J.TATE.2022.103802>
- Woolley, G. (2011). *Reading Comprehension*. <https://doi.org/10.1007/978-94-007-1174-7>
- Yildiz, Y., & Celik, B. (2020). The use of scaffolding techniques in language learning: Extending the level of understanding. *International Journal of Social Sciences & Educational Studies*, 7(3), 148–153. <https://doi.org/10.23918/ijsses.v7i3p148>
- Yow, W. Q., Tan, J. S. H., & Flynn, S. (2018). Code-switching as a marker of linguistic competence in bilingual children. *Bilingualism*, 21(5), 1075–1090. <https://doi.org/10.1017/S1366728917000335>
- Zairul, M. (2025). Mastering thematic analysis: a step-by-step guide for beginners with tips for systematic analysis using ATLAS.ti 25. *International Journal of Qualitative Methods*, 24. <https://doi.org/10.1177/16094069251384401>

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright ©2025 Jati Suryanto, Berliana Farras Rachmawati, Saefurrohman, Maryam Sorohiti. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.