p-ISSN 2252-7427, e-ISSN 2541-3635
https://ieflj.uniku.ac.id/pub

p-ISSN 2252-7427 e-ISSN 2541-3635
Indonesian EFL Journal
AISEE
The Association of Indonesia Scholars of English Education

IMPLEMENTATION OF DEEP LEARNING OF CRITICAL LISTENING SKILLS IN HIGHER EDUCATION: A STRATEGY TO IMPROVE THE QUALITY OF EDUCATION

Deby Luriawati Naryatmojo[1], Santi Pratiwi Tri Utami[2], Dyah Prabaningrum[3]
[1,2,3]*Indonesian Language and Literature Education, Universitas Negeri Semarang, Indonesia*

Corresponding author's email: debyluriawati@mail.unnes.ac.id

APA Citation: Naryatmojo, D. L., Utami, S. P. T., & Prabaningrum, D. (2026). Implementation of deep learning of critical listening skills in higher education: A strategy to improve the quality of education. *Indonesian EFL Journal, 12*(1), 257-266. https://doi.org/10.25134/ieflj.v12i1.86

Received: 14-09-2025 Accepted: 21-11-2025 Published: 30-01-2026

Abstract: This study aimed to evaluate the effectiveness of a deep learning-based instructional model in improving critical listening skills in higher education. A quantitative approach was employed using a quasi-experimental design with a nonequivalent control group. Two classes from Indonesian Language and Literature Education programs at two state universities with "A" accreditation were purposively selected. The instrument was a seven-item essay test that had been validated for reliability, item difficulty, and discrimination power. Assumption checks (Shapiro–Wilk and Levene tests) confirmed data normality and homogeneity of variance. Baseline differences were analyzed using independent-samples t-tests; within-group gains were tested with paired-samples t-tests; between-group posttest differences with independent-samples t-tests; and learning efficiency was summarized using the N-gain index. Both groups demonstrated significant improvement from the pretest to the posttest ($p < 0.001$). The experimental group achieved a higher mean posttest score of 83.54 ± 7.33 compared to the control group, which was 79.63 ± 5.31, representing a statistically significant difference ($p = 0.012$). The experimental group also demonstrated a significantly higher N-gain ($p = 0.020$), with a high-gain category (13.51%) observed only in this group. Deep learning-based instruction proved more effective than conventional teaching in enhancing critical listening skills, both in absolute achievement and learning efficiency. These findings support the adoption of deep learning pedagogy as an instructional strategy to strengthen students' critical literacy and 21st-century competencies in higher education.

Keywords: *critical listening; deep learning; higher education.*

INTRODUCTION

The imperative to enhance higher education quality is central to achieving Sustainable Development Goal 4, which advocates for inclusive, equitable, and quality education for all (Susanti & Chanifudin, 2025). Within this framework, the development of critical 21st-century competencies, particularly critical listening skills, is paramount for equipping students to navigate complex information landscapes, engage in reasoned discourse, and contribute to societal advancement (Gunawan, 2023; Baki, 2025; Costa & Cipolla, 2025). Critical listening transcends basic comprehension, involving the analytical evaluation of spoken discourse to discern meaning, assess credibility, and infer intent (Gülten & Zekerya, 2019; Schieble et al., 2021; Azzahra et al., 2023). Consequently, identifying pedagogical strategies that effectively cultivate these sophisticated cognitive abilities represents a critical focus for contemporary educational research, positioning deep learning as a promising instructional paradigm (Lin et al., 2023; Naseer et al., 2024; Ouahi et al., 2024).

A significant body of research underscores a strong positive correlation between critical listening proficiency and blended learning environments. Blended models that integrate online and face-to-face components have been shown to provide diverse auditory inputs and scaffolded activities, fostering improved

analytical listening (Yang et al., 2012; Caruso et al., 2017). Recent studies further confirm that structured, technology-mediated blended approaches can optimize the cognitive processes required for critical listening (Sujatha & Rajasekaran, 2024; Khallaf et al., 2024). The transition to digital and e-learning ecosystems in higher education has catalyzed the exploration of advanced computational tools, with deep learning algorithms offering unprecedented solutions for personalizing instruction, analyzing learner behavior, and providing adaptive feedback without necessitating extensive additional human resources (Dias et al., 2020; Muniasamy & Alasiry, 2020; Ginzburg et al., 2022).

Deep learning's application in education extends to creating flexible, student-centered learning pathways that adapt to individual needs, thereby enhancing overall educational quality and accessibility (Lin et al., 2023; Naseer et al., 2024). A particularly engaging application of this technology is gamification, which integrates game design elements into learning contexts to boost motivation, engagement, and commitment (Aguiar-Castillo et al., 2020; Khakpour & Colomo-Palacios, 2020). Empirical evidence indicates that gamified deep learning environments can significantly improve student attitudes and sustained focus, which are essential precursors to developing higher-order skills like critical listening (Aguiar-Castillo et al., 2019; Covrig et al., 2023; TATLI et al., 2023; Hafez et al., 2023).

Pedagogically, deep learning aligns with project-based and constructive learning approaches that are instrumental for fostering critical thinking and problem-solving (Sun & Ge, 2020; Wahyudiati et al., 2022). These active learning strategies facilitate deep information processing, collaborative inquiry, and reflective practice, which are crucial for mastering critical listening (Moustafa & Al-Rashaida, 2024; Purwati et al., 2024). The theoretical foundation is supported by social constructivism, which posits that knowledge is constructed through social interaction and reflection, a process that deep learning strategies can effectively mediate (Silalahi et al., 2022; Naryatmojo, 2018). Within the Indonesian context, research on listening skills has incorporated constructive models and local wisdom to foster national identity and cultural preservation, highlighting the need for critical thinking in information processing (Naryatmojo, 2019a; 2019b).

However, initial observations reveal a gap between potential and practice. The implementation of deep learning strategies for critical listening often remains suboptimal, failing to reach the highest pedagogical standards and learning objectives (D. L. Naryatmojo et al., 2024). This gap is partly attributable to challenges in selecting appropriate technological instruments and integrating them seamlessly into the curriculum (D. Susanti & Hernawati, 2020; Fu, 2024). Moreover, while technology provides tools, the core of critical listening involves a deep-thinking approach to draw informed conclusions, a skill that must be intentionally cultivated (Kazu & Demirlp, 2017; Oktariani & Ekadiansyah, 2020; Gunawan et al., 2023). Therefore, learning innovation that strategically leverages deep learning is required to maximize educational outcomes in higher education.

The relevance of this focus extends beyond academia. Enhancing critical listening skills prepares students to address complex socio-economic challenges, fostering social awareness and empowering them to act as agents of change in line with broader sustainable development goals (Gunawan, 2023; Costa & Cipolla, 2025). A successful implementation thus depends not only on technological infrastructure but also on organizational readiness and pedagogical design (Susanti & Hernawati, 2020; Ouahi et al., 2024). By addressing these multifaceted factors, deep learning can contribute substantively to developing a more adaptive, critical, and socially responsible graduate cohort, thereby supporting the creation of a sustainable society.

Despite growing interest in deep learning and critical listening separately, there is a paucity of empirical, quantitative research that directly tests the efficacy of a structured deep learning-based instructional model specifically designed to enhance critical listening skills in higher education. Existing literature extensively discusses theoretical potentials, technological applications, and correlated benefits (e.g., Aguiar-Castillo et al., 2020; Lin et al., 2023; Naseer et al., 2024), but lacks controlled experimental studies measuring its comparative effectiveness against conventional methods in a real classroom setting. This study, therefore, seeks to answer the following research question: How effective is a deep learning-based instructional model in improving the critical listening skills of university students compared to conventional teaching methods?

This study introduces novelty by integrating the conceptual framework of deep learning pedagogy—emphasizing data-driven reflection, metacognitive engagement, and adaptive learning—into a structured instructional model explicitly for critical listening development. It moves beyond theoretical proposition or technological demonstration to provide empirical, quasi-experimental evidence of the model's impact on both absolute performance and normalized learning gain. Furthermore, it contextualizes this innovation within the Indonesian higher education landscape, specifically targeting English-major students, thereby contributing unique insights from a Global South perspective on implementing advanced pedagogical technologies to cultivate essential 21st-century literacy skills.

The findings of this study hold significant theoretical, pedagogical, and practical implications. Theoretically, it strengthens the social constructivist framework by demonstrating how deep learning principles can operationalize active knowledge construction in listening comprehension (Silalahi et al., 2022; Naryatmojo, 2018). Pedagogically, it provides educators and curriculum designers with an evidence-based model for transforming passive listening instruction into an active, critical, and engaging cognitive process. Practically, it offers higher education institutions a viable strategy to enhance educational quality and graduate competencies, aligning with national and global goals for sustainable development (Susanti & Chanifudin, 2025; Ouahi et al., 2024). Ultimately, by demonstrating a pathway to more effective listening skill development, this research contributes to preparing a generation of learners equipped with the critical literacy necessary for personal, professional, and societal advancement.

METHOD

This study adopted a quantitative research paradigm, anchored in a post-positivist philosophical stance (Aguinis, 2023; Gamage, 2025). Post-positivism operates on the principle that reality exists but can only be approximated, making controlled measurement and objective analysis essential for testing hypotheses and building knowledge (Dubey & Kothari, 2022; Singh, 2022). Aligned with this stance, a quasi-experimental design was selected as the primary research strategy. Specifically, a nonequivalent control group design was implemented (Creswell, 2021; Mishra & Alok, 2022). This design is appropriate in educational settings where random assignment of intact classes is often logistically or ethically impractical, yet the comparison of different instructional interventions is necessary (Creswell & Plano Clark, 2023). The design involved pre-test and post-test measurements for both an experimental group, which received the deep learning-based intervention, and a control group, which received conventional instruction (Saharan et al., 2024; Zou & Xu, 2023).

The target population for this study was defined as undergraduate students enrolled in Indonesian Language and Literature Education programs at public universities in Indonesia with a high accreditation standing (Willie, 2024). The accessible population was further narrowed to students from two specific state universities: Universitas Negeri Semarang (UNNES) and Universitas Jenderal Soedirman (Unsoed) (Islam, Khan, & Baikady, 2022). Purposive sampling was employed to select the participating classes based on specific, pre-determined criteria to enhance comparability and control for extraneous variables (Gupta & Gupta, 2022; Willie, 2024). The criteria were: (1) the study program must hold an "A" accreditation rating, (2) the classes must be regular (non-international) sessions, and (3) the host institution must be a Public University with Legal Entity status (*PTN-BH*). One intact class from UNNES was assigned as the experimental group, and one intact class from Unsoed was assigned as the control group**.**

The procedure followed a structured sequence over an academic semester. First, both groups completed an identical pre-test to establish a baseline for critical listening skills. Following this, the intervention phase commenced. The experimental group received instruction via a deep learning-based model**.** This model was characterized by strategies promoting higher-order cognitive engagement, such as structured reflective journals on listening content, collaborative deconstruction of complex audio arguments, metacognitive exercises for monitoring listening comprehension, and project-based tasks requiring synthesis and evaluation of auditory information (Pal, 2023). The control group received conventional instruction, which followed a standard lecture-and-discussion format focused on content delivery and basic comprehension checks. Crucially, both groups covered the same core curriculum content, had equivalent total contact hours, and were taught by

instructors with comparable qualifications to ensure internal validity (Aguinis, 2023; Sreekumar, 2023). At the conclusion of the instructional period, both groups took an identical post-test to measure learning outcomes.

The primary instrument for data collection was a teacher-constructed critical listening essay test consisting of seven items. The test was designed to measure competencies across literal and interpretive listening domains (Mishra & Alok, 2022). To ensure validity and reliability, a multi-stage validation process was undertaken. First, content validity was established through expert judgment by three senior lecturers in language education. Subsequently, the test was piloted on a separate cohort of 26 students from the target population. Data from the pilot were analyzed for item validity using Pearson's correlation (item-total correlation), reliability using Cronbach's Alpha ($\alpha \geq 0.70$ considered acceptable), item difficulty index, and discrimination power (Javali & Javali, 2024; Saharan et al., 2024). All seven items met the established criteria for validity ($r >$ r-table, $p < 0.05$), reliability ($\alpha = 0.886$), appropriate difficulty (moderate), and adequate discrimination power, confirming the instrument's fitness for use.

Quantitative data analysis was conducted using SPSS software (Version 24). The analysis proceeded in distinct phases, adhering to established methodological protocols (Dehalwar & Sharma, 2023; Creswell & Inoue, 2025). First, classical assumption tests were performed. The normality of score distributions for both pre-test and post-test data was assessed using the Shapiro-Wilk test, and the homogeneity of variances between groups was tested using Levene's test (Gamage, 2025).

Following confirmation of parametric assumptions, a sequential statistical analysis was conducted. First, an independent samples t-test confirmed baseline equivalence between the groups via pre-test scores. Next, paired samples t-tests evaluated within-group improvements from pre-test to post-test for both groups. To compare intervention effectiveness, an independent samples t-test was then performed on the post-test scores. Finally, learning efficiency was analyzed using Normalized Gain (N-gain), calculated per student with the Hake formula [g = (post-test – pre-test)/(100 – pre-test)] and categorized as low (<0.30), medium (0.30–0.70), or high (≥0.70). An independent samples t-test was also applied to the mean N-gain scores to compare the magnitude of improvement between the experimental and control groups.

This study was conducted in accordance with standard ethical principles in social science research (Denscombe, 2025; Islam, Khan, & Baikady, 2022). Informed consent was obtained from all participating students and instructors. Confidentiality and anonymity of participants' data were strictly maintained. To enhance methodological rigor, several strategies were employed: triangulation of data sources (pre-test, post-test, gain scores), use of a validated and reliable instrument, control for key confounding variables through purposive sampling, and adherence to a transparent, pre-planned analytical framework (Aguinis, 2023; Creswell, 2024). This structured approach ensured the credibility, dependability, and confirmability of the research findings.

RESULTS AND DISCUSSION

The effectiveness of the deep learning intervention was assessed by comparing pre-test and post-test results from both the experimental and control classes. Prior to hypothesis testing, the data were checked to ensure they met the necessary statistical assumptions for parametric analysis. Normality tests (Kolmogorov-Smirnov) confirmed that the score distributions for all tests (pre- and post-test in both groups) did not significantly deviate from a normal distribution (all $p > 0.05$). Additionally, Levene's tests for equality of variances indicated no significant differences in score variability between the two groups for either the pre-test or post-test (all $p > 0.05$). These results, summarized in Table 4, confirmed that the assumptions of normality and homogeneity of variance were met, validating the use of subsequent parametric statistical tests.

Table 4. *Results of the assumption checks for pre- and post-test scores in the control and experimental classes.*

Evaluation	Treatment Groups	One-sample Kolmogorov-Smirnov		Levene's test	
		Sig.	Description	Sig.	Description
Pre	Control	0.067	normal	0.616	homogen

	Experiment	0.200	normal		
Post	Control	0.113	normal	0.061	homogen
	Experiment	0.124	normal		

Note. For assumption checks, *a p-value > 0.05 indicates no evidence against normality/ homogeneity (the* assumptions are met). "Sig." = *p*-value.

With the assumptions of normality and homogeneity of variance satisfied, parametric t-tests were deemed appropriate for data analysis. A preliminary independent samples t-test confirmed no significant difference in pre-test scores between the experimental and control groups ($p = 0.800$), establishing baseline equivalence in listening comprehension. This initial parity is crucial for internal validity, as it ensures that any post-intervention differences are more likely attributable to the treatment rather than pre-existing ability gaps (Creswell & Plano Clark, 2023).

Following the intervention, a comparison of post-test scores revealed a statistically significant advantage for the experimental group that received deep learning-based instruction ($p < 0.05$). This finding indicates that the pedagogical model, which incorporated strategies for metacognitive engagement and reflective practice, effectively enhanced students' critical listening abilities beyond the gains achieved through conventional instruction. The result aligns with theoretical perspectives on deep learning, which posit that such approaches foster advanced cognitive processing and a more analytical stance toward information (Silalahi et al., 2022; Pal, 2023).

The significant improvement observed corroborates prior research highlighting the transformative potential of student-centered, technology-mediated pedagogies in language education. Studies suggest that when learners are engaged in deconstructing complex inputs and are encouraged to apply knowledge reflectively, they develop stronger competencies in interpreting meaning and evaluating spoken discourse (Lin et al., 2023; Gunawan et al., 2023). This shift from mechanical comprehension to critical engagement supports broader educational objectives focused on cultivating analytical reasoning and self-regulated learning as essential 21st-century skills (Baki, 2025; Costa & Cipolla, 2025).

In summary, the confirmation of baseline group equivalence, coupled with the significant post-test performance difference, robustly demonstrates the efficacy of the deep learning-based instructional model. These findings not only validate the intervention but also suggest its potential for wider application in higher education to advance critical literacy. Future research could explore longitudinal effects and the integration of these strategies within interdisciplinary digital learning environments to further solidify and expand upon these pedagogical benefits.

Following the confirmation of parametric assumptions, statistical analyses proceeded. **Table 5** presents the key comparative results. An independent samples t-test confirmed that before the intervention, the experimental ($M = 70.57$, $SD = 6.07$) and control ($M = 69.31$, $SD = 7.32$) groups did not differ significantly in pre-test scores ($p = .800$), establishing crucial baseline equivalence. This procedural rigor ensures that subsequent differences can be attributed to the treatment, strengthening the study's internal validity—a fundamental principle in quasi-experimental design (Creswell & Plano Clark, 2023).

The post-intervention data in **Table 5** reveal the core finding. While both groups showed significant within-group improvement ($p < .001$), the experimental group achieved a substantially higher mean post-test score (83.54 ± 7.33) compared to the control group (79.63 ± 5.31). An independent samples t-test confirmed this between-group difference was statistically significant ($p = .012$). This result provides direct empirical support for the hypothesis, demonstrating that the deep learning-based model was more effective than conventional instruction.

This superior performance aligns with and extends previous research on technology-enhanced, student-centered pedagogies. The significant gain suggests that the model's emphasis on reflective practice, collaborative problem-solving, and metacognitive engagement (Pal, 2023) successfully fostered the advanced cognitive processing required for critical listening. Such an outcome corroborates studies which assert that deep learning strategies move students beyond surface-level comprehension toward the interpretative and evaluative skills essential for critical literacy (Silalahi et al., 2022; Lin et al., 2023).

The data summarized in **Table 5**—showing equivalent starting points but diverging outcomes—offers compelling evidence for the

efficacy of the deep learning intervention. These findings validate the pedagogical model and contribute to the growing consensus on the need for innovative, cognitively engaging frameworks in higher education to develop foundational 21st-century competencies.

Table 5. *Significance of mean differences in pre- and post-test scores in the control and experimental groups*

Evaluation	Control	Experimental	p-value
Pre-test	69.31 ± 7.32	70.57 ± 6.07	0.800
Post-test	79.63 ± 5.31	83.54 ± 7.33	0.012
p-value	0.000	0.000	

Note. Significance was evaluated at $\alpha = 0.05$ (95% confidence level) using one-tailed independent-samples t-tests; results with $p \leq 0.050$ were considered significant.

Following the instructional intervention, both groups demonstrated notable progress in their post-test performance. However, the experimental class attained a superior mean score (83.54 ± 7.33) compared to the control class (79.63 ± 5.31). The between-group difference was statistically significant according to a one-tailed independent-samples *t*-test ($p = 0.012$), suggesting that the instructional model implemented in the experimental class was more effective than the conventional approach. Moreover, within-group comparisons from pre-test to post-test also revealed significant improvements (paired-samples tests, both $p < 0.001$), confirming that substantial learning gains occurred in both groups. Collectively, these findings affirm the positive influence of the intervention, with the experimental group experiencing the most pronounced improvement, consistent with the higher percentage increase and *n*-gain presented in Figure 1.

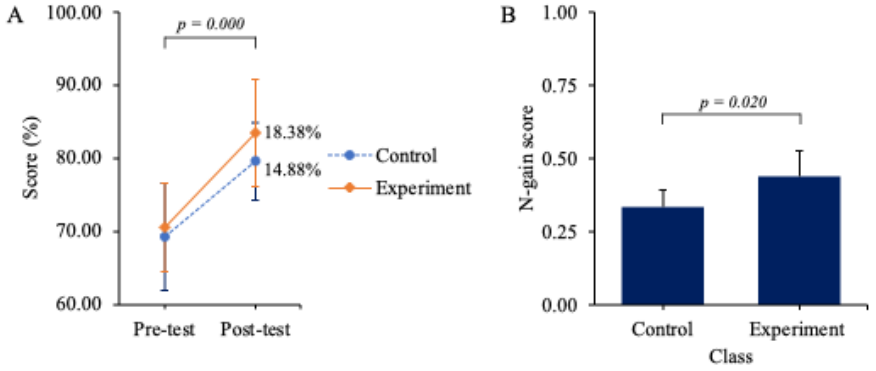


Figure 1. *Significant differences in percentage improvement (A) and normalized gain (n-gain) (B) between the control and experimental groups.*

N-gain categories (low: $g < 0.30$; medium: $0.30 \leq g < 0.70$; high: $g \geq 0.70$). N-gain was computed as (post-test − pre-test) / (100 − pre-test), and the percentages were calculated as (count/N) × 100%. Statistical significance was assessed using one-tailed independent-samples t-tests at the 95% confidence level ($\alpha = 0.05$); results with $p \leq 0.050$ were considered significant.

As shown in Panel A, both groups experienced increases in scores from pre-test to post-test, with improvement rates of 18.38% for the control group and 14.88% for the experimental group. Although both improvements were statistically significant ($p = 0.000$), the higher post-test mean achieved by the experimental class indicates the superior efficacy of the instructional intervention compared to the traditional teaching method. This conclusion is further reinforced by the *n*-gain analysis (Figure 1B), which measures learning efficiency relative to the maximum attainable improvement. The experimental group's mean *n*-gain exceeded that of the control group, and the difference was statistically significant ($p = 0.020$). These findings suggest that the experimental group's instructional approach not only improved learning outcomes in absolute terms but also enhanced the efficiency of knowledge acquisition.

The statistically significant post-test difference between groups ($p = 0.012$) supports the greater effectiveness of the deep-learning–based instructional model compared to the conventional teaching framework. This outcome suggests that

learners in the experimental class, who engaged in data-driven, reflective, and metacognitive learning processes, developed more durable and transferable listening comprehension skills. In contrast, improvements in the control group may broadly reflect surface-level practice effects rather than more profound cognitive restructuring. These findings align with theories of meaningful learning, which propose that learners who actively construct and internalize knowledge through iterative cognitive engagement exhibit more substantial and enduring gains in performance.

Furthermore, the consistently significant within-group differences ($p < 0.001$) demonstrate that both instructional approaches successfully facilitated measurable learning progress. Nevertheless, the magnitude of improvement in the experimental group underscores the additional value of incorporating deep-learning principles such as pattern recognition, multimodal input processing, and self-directed problem-solving into listening instruction. These cognitive mechanisms likely encouraged learners to progress beyond passive comprehension toward analytical interpretation and inferential reasoning, both of which are essential for advanced listening competence in academic and professional domains.

The observed increase in mean scores across both groups also reflects the motivational and cognitive benefits of structured instructional design. However, the experimental group's superiority in both absolute improvement and *n*-gain efficiency indicates a more accelerated and cognitively optimized learning trajectory. From an educational psychology perspective, such a pattern is often interpreted as evidence of enhanced *transfer of learning*, whereby students can apply previously acquired strategies to novel listening contexts. This is particularly significant in higher education, where the objective extends beyond mere comprehension accuracy toward cultivating autonomous, critical, and adaptive listeners capable of engaging with complex linguistic input.

Moreover, the higher *n*-gain value achieved by the experimental group reinforces the argument that deep-learning–oriented instruction enhances knowledge acquisition efficiency relative to learners' initial proficiency levels. Because the *n*-gain metric accounts for potential ceiling effects, it provides a more nuanced understanding of learning improvement. The statistically significant *n*-gain difference ($p = 0.020$) implies that the deep-learning framework not only elevated overall performance but also systematically strengthened learners' ability to process complex auditory information and derive meaning from contextually embedded cues. These results highlight the pedagogical relevance of integrating computational thinking and reflective metacognition into listening instruction.

In summary, the data presented in Table 5 and Figure 1 collectively provide convergent evidence that the deep-learning–based instructional intervention significantly enhanced both the magnitude and efficiency of learning outcomes in listening comprehension. Beyond mere statistical significance, the pedagogical implications of these findings are profound: embedding deep-learning strategies within language instruction can transform passive listening exercises into cognitively active and engaging learning experiences. Such transformation aligns with contemporary educational paradigms emphasizing learner autonomy, technological mediation, and the development of higher-order cognitive skills as fundamental components of 21st-century language education. These findings thus provide a robust empirical foundation for subsequent discussions regarding the theoretical and pedagogical implications of implementing deep-learning principles in practical classroom settings.

Subsequently, individual-level improvement classifications are summarized in Table 6.

Table 6. *Frequency distribution of students across n-gain categories*

N-gain Category	Control		Experiment	
	N	%	N	%
Low	13	37.14	11	29.73
Moderate	22	62.86	21	56.76
High	-	-	5	13.51

In the control group, most students fell into the moderate improvement category (22 students, 62.86%), while a sizeable portion still showed low improvement (13 students, 37.14%). Importantly, no student in the control group reached the high n-gain category, indicating that the conventional teaching method improved learning, but not to an optimal level. In the

experimental group, the pattern is more favorable. A majority of students were also in the moderate category (21 students, 56.76%), but the proportion of low improvers was smaller than in the control group (29.73% vs 37.14%). Most notably, five students in the experimental group (13.51%) achieved a high n-gain, whereas none in the control group did. The presence of a “high-gain” segment only in the experimental class suggests that the deep–learning–based instruction was able to elevate some students to a level of learning improvement that the conventional method could not achieve.

A deep learning approach can significantly improve the students' critical thinking skills in project-based learning. Deep learning plays a crucial role in fostering a deeper conceptual understanding through active, collaborative, and reflective learning activities. In-depth information processing, including reflective discussions and collaborative problem-solving, can enhance students' listening and critical thinking skills. They emphasize the importance of active students’ involvement to construct meaning through contextual and meaningful interactions. Furthermore, deep learning-based models are effective in the learning process, enabling students to analyze abstract materials more effectively, particularly in language skills such as listening and writing. This type of learning enables students to connect new knowledge with prior experiences, making it more meaningful.

The application of deep learning also aligns with Vygotsky's theory of social constructivism, which argues that learning is more effective when students engage in social activities, discussions, and reflection (Silalahi et al., 2022). The students’ engagement fosters more profound and more lasting understanding. The deep learning-based learning strategies significantly improve the students' critical listening skills. Therefore, this approach is highly suitable for implementation in higher education as part of efforts to improve critical literacy and 21st-century competencies.

CONCLUSION

This study demonstrates that instruction grounded in deep learning principles significantly improves students' critical listening skills over conventional teaching methods. The established baseline equivalence between groups confirms that the superior post-test performance of the experimental group is a direct result of the pedagogical intervention. This superiority is evidenced by higher mean scores, greater percentage improvement, and, most tellingly, a significantly higher normalized gain (N-gain). The fact that students achieving a "high-gain" classification were found only in the experimental cohort underscores the unique potential of this approach to facilitate transformative learning.

These findings validate theoretical frameworks that prioritize active cognitive engagement, reflective practice, and collaborative problem-solving as mechanisms for meaningful learning. The deep learning model successfully guides learners past rote comprehension toward deeper interpretive and evaluative processes. This aligns with social constructivist views, where knowledge is built through interaction and reflection, leading to more durable and transferable understanding.

In summary, deep learning-based instruction offers a demonstrably more effective and efficient route to developing advanced listening comprehension and critical thinking. Its adoption represents a strategic and relevant response to the pressing need in higher education to foster the autonomous, analytical competencies required for the 21st century.

REFERENCES

Abidin, M. Z., & Halis, F. D. (2024). Pengembangan aplikasi online berbasis gamifikasi untuk meningkatkan kompetensi bahasa Inggris listening dan reading comprehension dalam persiapan Toeic test. *Jurnal Teknologi Pembelajaran Indonesia*, *14*(2). https://doi.org/10.23887/jurnal_tp.v14i2.4047

Aguiar-Castillo, L., Clavijo-Rodríguez, A., López, L. H., Saa-Pérez, P. D., & Pérez-Jiménez, R. (2020). Gamification and deep learning approaches in higher education. *Journal of Hospitality, Leisure, Sport & Tourism Education*, *29*, Artikel 100290. https://doi.org/10.1016/j.jhlste.2020.100290

Aguiar-Castillo, L., Clavijo-Rodríguez, A., Saa-Pérez, P. De, Pérez, R., & Jiménez. (2019). Gamification as an approach to promote tourist recycling behavior. *Sustainability*, *11*(8). https://doi.org/10.3390/su11082201

Azzahra, A., Ibrahim, H., Rohimah, N., Nasution, S. F., & Zakiyyah, Z. (2023). Mengembangkan keterampilan menyimak yang kritis di kelas tinggi. *CENDIKIA: Jurnal Ilmu Sosial, Bahasa, dan Pembelajarannya*, *3*(3). https://doi.org/10.55606/cendikia.v3i3.1380

Baki, Y. (2025). The impact of critical listening and critical reading on critical thinking. *Behavioral Sciences*, *15*(1), 34. https://doi.org/10.3390/bs15010034

Caruso, M., Gadd, A., & Tebbit, S. (2017). Teaching how to listen: Blended learning for the development and assessment of listening skills in a second language. *Journal of University Teaching and Learning Practice*, *14*(1). https://doi.org/10.53761/1.14.1.7

Costa, M. F. B., & Cipolla, C. (2025). Critical soft skills for sustainability in higher education: A multi-phase qualitative study. *Sustainability*, *17*(2), 377. https://doi.org/10.3390/su17020377

Covrig, M., Agoston, S., Igreț, R. Ș., Marinaș, C. V., Miron, A. D., & Roman, M. (2023). Students' engagement and motivation in gamified learning. *Amfiteatru Economic*, *25*, 1003. https://doi.org/10.24818/ea/2023/s17/1003

Dias, S. B., Hadjileontiadou, S. J., Diniz, J. A., & Hadjileontiadis, L. J. (2020). DeepLMS: A deep learning predictive model for supporting online learning in the Covid-19 era. *Scientific Reports*, *10*(1). https://doi.org/10.1038/s41598-020-76740-9

Fu, L. (2024). The design of a teaching system for a university English digital course based on a deep learning model. *Applied Mathematics and Nonlinear Sciences*, *9*(1). https://doi.org/10.2478/amns-2024-0204

Ginzburg, O. (2021). *The nature of giant clumps in high-z discs: A deep-learning comparison of simulations and observations* [Preprint]. arXiv. https://arxiv.org/pdf/2011.06616

Ginzburg, S., Fuller, J., Kawka, A., & Caiazzo, I. (2022). Slow convection and fast rotation in crystallization-driven white dwarf dynamos. *Monthly Notices of the Royal Astronomical Society*, *514*(3), 4111–4119. https://doi.org/10.1093/mnras/stac1363

Gülten, E., & Zekerya, B. (2019). Activity suggestions for improving critical listening skills. *Educational Research and Reviews*, *14*(17), 639. https://doi.org/10.5897/err2019.3808

Gunawan, A., Gustine, G. G., & Gunawan, M. H. (2023). Critical listening in higher education: Insights from Indonesian EFL learners. *English Review: Journal of English Education*, *11*(3), 675. https://doi.org/10.25134/erjee.v11i3.8236

Hafez, D. W. A., Salam, M. A. E., Farid, S., & Farouk, M. (2023). The impact of gamification on intention to use in e-learning through student attitude: Evidence from Egyptian private higher education institutions (HIEs). *SocioEconomic Challenges*, *7*(2), 67–77. https://doi.org/10.21272/sec.7(2).67-77.2023

Karim, S. A., Sudiro, S., & Sakinah, S. (2021). Utilizing test items analysis to examine the level of difficulty and discriminating power in a teacher-made test. *EduLite: Journal of English Education, Literature and Culture*, *6*(2), 256.

Kazu, H., & Demiralp, D. (2017). Comparison of critical listening proficiency of teacher candidates in terms of several variables. *Eurasian Journal of Educational Research*. https://dergipark.org.tr/tr/pub/ejer/issue/42457/511293

Khakpour, A., & Colomo-Palacios, R. (2020). Convergence of gamification and machine learning: A systematic literature review. *Technology, Knowledge and Learning*, *26*(3), 597. https://doi.org/10.1007/s10758-020-09456-4

Khallaf, M. N., R.-B., A. A., M., & Abugoha. (2024). Genre analysis in blended learning: Empowering English majors' critical listening skills: Case of Faculty of Education, Minia University, Egypt. *Forum for Linguistic Studies*, *6*(5).

Kwang Gi Kim, D. (2016). Deep learning. *Healthcare Informatics Research*, *22*(4), 351. https://doi.org/10.4258/hir.2016.22.4.351

Lin, Y., Chen, H., Xia, W., Fan, L., Wu, P., Wang, Z., & Li, Y. (2023). *A comprehensive survey on deep learning techniques in educational data mining* [Preprint]. arXiv. https://doi.org/10.48550/arxiv.2309.04761

Moustafa, A., & Al-Rashaida, M. (2024). Fostering students' critical thinking through the implementation of project-based learning. In *Advances in Educational Technologies and Instructional Design* (p. 42). IGI Global. https://doi.org/10.4018/979-8-3693-0880-6.ch004

Muniasamy, A., & Alasiry, A. (2020). Deep learning: The impact on future eLearning. *International Journal of Emerging Technologies in Learning (iJET)*, *15*(1), 188. https://doi.org/10.3991/ijet.v15i01.11435

Naryatmojo, D. L. (2018, Oktober). Penggunaan taksonomi bloom dalam pembelajaran keterampilan menyimak bermuatan pendidikan karakter profetik untuk mengukur keberhasilan hasil belajar mahasiswa. Dalam *Pertemuan Ilmiah Bahasa dan Sastra Indonesia (PIBSI)*.

Naryatmojo, D. L. (2019a). Internalization of the concept of local wisdom for students in the listening class. *Arab World English Journal (AWEJ)*, *10*(1).

Naryatmojo, D. L. (2019b). *Model lingkungan belajar konstruktif pada pembelajaran menyimak*. Cipta Prima Nusantara.

Naryatmojo, D. L., Haryadi, & Prasandha, D. (2024). The effectiveness of constructive learning environment models in learning listening skills to improve college student learning outcomes. *RETORIKA: Journal of Language, Literature, and Its Teaching*, *17*(1).

Naseer, F., Khan, M. N., Tahir, M., Addas, A., & Aejaz, S. M. H. (2024). Integrating deep learning techniques for personalized learning pathways in higher education. *Heliyon*, *10*(11), Artikel e32628. https://doi.org/10.1016/j.heliyon.2024.e32628

Ndraha, L., Halawa, N., Harefa, N. A. J., & Ndruru, M. (2023). Peningkatan kemampuan menyimak cerpen dengan media pembelajaran audiovisual peserta didik kelas VII SMP Negeri 1 Hiliserangkai tahun pembelajaran 2024/2025. *Jurnal Kata: Bahasa, Sastra, dan Pembelajarannya*, *13*(1), 292–300. http://dx.doi.org/10.23960/Kata

Oktariani, & Ekadiansyah, E. (2020). Peran literasi dalam pengembangan kemampuan berpikir kritis. *Jurnal Penelitian Pendidikan, Psikologi dan Kesehatan*, *1*(1), 23–33.

Ouahi, M., Khoulji, S., & Kerkeb, M. L. (2024). Analysis of deep learning development platforms and their applications in sustainable development within the education sector. *E3S Web of Conferences*, *477*, 98. https://doi.org/10.1051/e3sconf/202447700098

Purwati, T., Setyarini, S., Waly, M. M., & Sumanto, E. (2024). Project-based learning integrated with Higher-Order Thinking Skills (HOTS) assessment to enhance the students' speaking skills in the EFL classroom. *Issues in Language Studies*, *13*(1), 322. https://doi.org/10.33736/ils.6090.2024

Schieble, M., Vetter, A., & Martin, K. M. (2021). Critical listening for critical conversations. *The English Journal*, *111*(2), 73. https://doi.org/10.58680/ej202131501

Silalahi, S. A., Zainal, A., & Sagala, G. H. (2022). The importance of deep learning on the constructivist approach. *Proceedings of the 2nd International Conference of Strategic Issues on Economics, Business, and Education (ICoSIEBE 2021)*, *204*, 243–246. https://doi.org/10.2991/aebmr.k.220104.036

Sugiyono. (2007). *Metode penelitian pendidikan: Pendekatan kuantitatif, kualitatif, dan R&D*. Alfabeta.

Sujatha, U., & Rajasekaran, V. (2024). Optimising listening skills: Analysing the effectiveness of a blended model with a top-down approach through cognitive load theory. *MethodsX*, *12*, 102630. https://doi.org/10.1016/j.mex.2024.102630

Sun, Q., & Ge, Z. (2020). A survey on deep learning for data-driven soft sensors. *IEEE Transactions on Industrial Informatics*. https://doi.org/10.1109/TII.2021.3053128

Susanti, D., & Hernawati, C. (2020). Keterampilan berpikir kritis mahasiswa dalam menyelesaikan soal Higher Order Thinking Skills. *EDU-MAT: Jurnal Pendidikan Matematika*. https://ppjp.ulm.ac.id/journal/index.php/edumat

Susanti, E., & Chanifudin. (2025). Upaya menuju pendidikan yang berkualitas dalam mewujudkan SDGS. *INNOVATIVE: Journal of Social Science Research*, *5*(1). https://j-innovative.org/index.php/Innovative

Tatli, Z., Gülay, A., & Mert, A. (2023). Impact of gamification applications on students' attitudes towards lesson and procrastination behaviors. *International Journal of Contemporary Educational Research*, *10*(2), 522. https://doi.org/10.52380/ijcer.2023.10.2.387

Wahyudiati, D., Irwanto, I., & Ningrat, H. K. (2022). Improving pre-service chemistry teachers' critical thinking and problem-solving skills using project-based learning. *World Journal on Educational Technology: Current Issues*, *14*(5), 1291. https://doi.org/10.18844/wjet.v14i5.7268

Yang, Y., Chuang, Y.-C., Li, L.-Y., & Tseng, S.-S. (2012). A blended learning environment for individualized English listening and speaking, integrating critical thinking. *Computers & Education*, *63*, 285. https://doi.org/10.1016/j.compedu.2012.12.012