EVALUATING THE IMPACT OF AI CHATBOTS ON STUDENT LEARNING OUTCOMES AT THE DIPLOMATIC ACADEMY OF VIETNAM

Lien Tran
Diplomatic Academy of Vietnam
Hai Nam Pham*
Diplomatic Academy of Vietnam
Do Ngan Giang Nguyen
Diplomatic Academy of Vietnam
Thanh Mai Le
Diplomatic Academy of Vietnam
Duc Minh Pham
Diplomatic Academy of Vietnam

Abstract: *This study rigorously assesses the impact of Artificial Intelligence (AI) integration, specifically through chatbots, on student learning outcomes at the Diplomatic Academy of Vietnam. Employing a mixed-methods approach, the research systematically investigates how these AI tools affect language proficiency, critical thinking abilities, and student engagement among 520 undergraduate students enrolled in diplomatic training courses across various disciplines, including international relations, economics, communication, and linguistics. The analysis is grounded in several educational theories: Constructivist Learning Theory underscores the interactive, student-centered learning experiences enabled by chatbots; the Technology Acceptance Model (TAM) explains their adoption based on perceived usefulness and ease of use; Self-Determination Theory (SDT) evaluates how chatbots influence students' motivation by supporting autonomy, competence, and relatedness; and Cognitive Load Theory assesses how chatbots help manage cognitive effort, enhancing comprehension and retention. Findings from pre- and post-tests reveal statistically significant improvements in student performance, while surveys and interviews highlight increased engagement and confidence in using AI-driven learning tools. However, challenges such as feedback accuracy limitations, technical constraints, and concerns about over-reliance on AI are also noted. This study provides evidence-based insights into the benefits and challenges of AI chatbots in diplomatic education, offering strategies for*

balancing AI-driven and traditional teaching methods to optimize learning. By addressing the gap in research on AI's role in specialized education, this paper contributes valuable recommendations for enhancing curriculum design and AI adoption in higher education, particularly in fields requiring critical thinking and professional communication skills.

Keywords: *AI chatbots in education, student learning outcomes, diplomatic training & AI*

INTRODUCTION

Background

AI's facilitation within the educational sector has substantially redefined conventional teaching and learning methods, providing advanced measures to improve students' results and overall performance (Holmes et al., 2019; Zawacki-Richter et al., 2019). Among these innovations are AI chatbots, widely considered among the most beneficial AI applications in education due to their capacity to provide instant feedback, create virtual learning environments, and offer individualized attention to learners (Adamopoulou & Moussiades, 2020; Okonkwo & Ade-Ibijola, 2021). These features address the growing need for flexible and adaptive educational approaches as the world continues to digitalize (Gokcearslan et al., 2024; Chang et al., 2022).

In education, higher learning systems, especially in the specialized areas such as diplomacy, AI chatbots can do more than just aid in specialized vocabulary. They can perform the functions of intelligent tutors who can assist students in dealing with difficult course content while developing necessary skills such as critical thinking and engagement as well. This is particularly important in institutions like the Diplomatic Academy of Vietnam where students are expected to master not only language but also possess strong analytical and interpersonal skills that are critical for their future careers as diplomats.

However, even though AI chatbots are becoming increasingly common, not enough research has been conducted on how these technologies

affect student learning outcomes in specialized areas. While studies in various countries emphasize the advantages of AI in education in general (Holmes et al., 2019), not much attention has been given to these specific issues of diplomatic education.This study seeks to bridge that gap by evaluating the effectiveness of AI chatbots in enhancing language proficiency, critical thinking, and engagement among students at the Diplomatic Academy of Vietnam.

Problem Statement

The Diplomatic Academy of Vietnam, as an education provider to future diplomats, has numerous unique factors that would wish to fulfill the requirements of a student in this excessively globalized, multilingual, and technology-centered world. However, modern teaching methods, although helpful, often do not emotionally connect and form a bridge with the new age students who are more open to personalized and engaging lessons. These challenges can be mitigated using AI chatbots, but so far, they have not been sufficiently studied in the scope of academic program integration.

Not much has been said in the existing literature on the effects of AI chatbots on critical academic outcomes such as speaking skills, active thinking, and participation in specialized areas like diplomacy. Holmes et al. (2019) showed how transformational AI technologies can be in the language learning environment, just like Chen et al. (2024) who pointed out the overwhelming potential of AI chatbots and other tools to serve as multi-role pedagogical agents, increasing engagement significantly in almost all educational activities. Still, those would need to be checked on target diplomacy education. In addition, the benefits and shortcomings of these devices in the context of diplomatic education has not been sufficiently discussed from students' as well as lecturers' point of view. Until one makes an evaluation on the degree of effectiveness it is very hard to realize the prospect AI chatbots available for this area of learning.

To mitigate these gaps, this research examines the effect of AI chatbots on the student learning outcomes at the Diplomatic Academy of Vietnam. Incorporating both quantitative and qualitative analysis ensures that the recommendations to implement AI chatbots into AI training programs are evidence-based and meet the needs of the institution as well as the students.

Research Questions

To achieve the aims of this study, the following research questions have been formulated:

1. How do AI chatbots impact students' language proficiency, critical thinking skills, and engagement at the Diplomatic Academy of Vietnam?

2. What are the key benefits and challenges of integrating AI chatbots into academic settings, as perceived by students?

Significance of the Study

The use of Artificial Intelligence (AI) in education, especially in specific areas like diplomacy, is elaborated by Okonkwo and Ade-Ibijola (2021). With the potential to transform personalized learning and engagement, the use of AI chatbots in education is profound. However, the research does not address its use in specialized training or diplomatic education which is a gap this study aims to fulfill. In doing so, this study contributes to the existing body of knowledge surrounding the incorporation of AI in higher education.

The provided insights are greatly useful for educators and policymakers, particularly at the Diplomatic Academy of Vietnam. Gokcearslan et al. (2024) recommended moderation of AI use in conjunction with other tools in order to solve new problems arising from cognitive overload or low-level tech acceptability. Building on this, the study proposes the incorporation of AI chatbots into the classroom to meet high educational standards and improve student learning.

Furthermore, the findings of the study highlight how crucial it is to predict and prepare for the changes that will come in the delivery of diplomatic education. Ait Baha et al. (2022) showed that chatbots have broad potential in improving students' critical thinking and analysis, skills highly needed in diplomacy to be able to interact with different stakeholders. Teachers can utilize AI-based solutions to fill in the gaps in traditional teaching methods in order to prepare future diplomats with appropriate levels of technological literacy and advanced AI-driven problem-solving skills.

Ultimately, these findings are important on a global scale, as they have significant consequences. Chang et al. (2022) studied the use of chatbots as a novel tool to solve learning complexities within and across various fields of study. This research is important because it demonstrates the potential of AI to improve learning in more advanced and specialized fields like diplomacy, and in doing so offers supporting evidence for the use of AI in education globally.

LITERATURE REVIEW

AI in Education

It is undebatable that the dynamics of learning have undergone an immense shift over the years, and the rise of Artificial Intelligence (AI) in education has been a perquisite to it. The first known form of Artificial Intelligence (AI) through the use of intelligent tutoring systems was recorded in the year 1980, which enabled learners to receive custom feedback alongside guided instructions to them. These systems made significant speak progress towards novel AI applications like chatbots, which presently possess state-of-the-art functions such as natural language processing, adaptive learning, etc.

AI has disrupted the field of education and particularly language learning, in various ways. AI tools offer an interactive and immersive environment in real time that helps learners in effectively acquiring a brand new language (Zawacki-Richter, 2019). For instance, learners

can use voice technology to practice their pronunciation, and AI driven writing tools assist in providing instantaneous corrections and suggestions. In the same manner, Al-Zahrani (2023) supported the use of AI in education by stating that AI tools are influential in problem buffering as students are developed to reason critically through the use of open-ended questions and are guided through questioning techniques.

The aforementioned advancements have not only enhanced the learners' mastery of language, but also facilitated the development of critical thinking skills, which is an important aspect in higher learning. Therefore AI based tools encourage learners to think critically and analytically by offering multifaceted, scenario based tasks which require the student to use theoretical knowledge in practical situations, according to Lawasi et al. (2024). Nonetheless, despite the usage of these specific tools in general education, their effectiveness in particular disciplines like diplomacy is still lacking, thus calling for further investigation of the phenomena.

AI Chatbots and Learning Outcomes

AI chatbots have shown significant potential to enhance learning outcomes, particularly in language proficiency, engagement, and critical thinking. Various studies have demonstrated their effectiveness in improving these aspects of education.

In the context of language proficiency, chatbots provide learners with real-time feedback and opportunities to practice conversational skills in a low-pressure environment. Hill et al. (2015) highlighted that chatbots enable learners to engage in meaningful interactions, fostering vocabulary acquisition and contextual language use. Similarly, Farrokhnia et al. (2023) emphasized the role of chatbots in improving speaking and writing fluency by providing adaptive language support and reducing performance anxiety.

Performance anxiety, especially in language learning, is a well-documented barrier that can hinder students' willingness to practice speaking and limit their language development. Recent research has explored how AI chatbots can specifically address this challenge by creating a supportive and low-pressure environment.

For example, Farrokhnia et al. (2023) conducted a study involving university students who used AI-powered chatbots for English language practice. Their findings revealed a statistically significant reduction in self-reported speaking anxiety after several weeks of chatbot interaction. Students attributed this improvement to the chatbot's nonjudgmental feedback and the ability to practice as often as needed without fear of negative evaluation from peers or instructors. Notably, the study found that students who reported higher initial anxiety experienced the greatest reductions, highlighting the chatbot's unique role as a "safe practice partner."

Similarly, Chang et al. (2022) examined the use of a mobile AI chatbot for nursing students learning English. Their mixed-methods research found that regular chatbot-based conversational practice not only improved language proficiency but also increased learner confidence and willingness to speak in class. Qualitative interviews indicated that students felt more comfortable experimenting with new vocabulary and complex sentence structures with the chatbots than in front of classmates, suggesting that chatbots help bridge the gap between passive knowledge and active usage.

Hill et al. (2015) compared human–human online conversations with human–chatbot interactions. Their study showed that learners were more willing to take conversational risks and made more attempts at complex sentence structures when using chatbots. The authors attributed this to the absence of social judgment and the patient, adaptive nature of the AI, which supported the gradual reduction of performance anxiety over time.

Engagement is another area where chatbots have proven highly effective. Adamopoulou and Moussiades (2020) emphasized that chatbots simulate human-like interactions, which create a sense of companionship and personalized attention for learners. This interactive element motivates students to participate actively in learning tasks. Research by Okonkwo and Ade-Ibijola (2021) further supports this, noting that chatbots increase learner engagement by delivering adaptive and immediate responses tailored to individual needs.

When it comes to fostering critical thinking, chatbots present learners with scenario-based tasks and open-ended questions. Lim and Makany (2023) demonstrated how generative AI-powered chatbots could effectively build students' critical thinking skills by presenting complex, reflective tasks. They also emphasized the need for careful deployment to ensure the purposeful integration of chatbots in higher education.

However, while these benefits are evident, the effectiveness of chatbots in specialized fields like diplomacy remains underexplored. This gap underscores the need for further research to understand how chatbots can support the development of domain-specific skills in complex disciplines.

Theoretical Frameworks

Constructivist learning theory

Constructivist Learning Theory emphasizes that knowledge is actively constructed by learners through experiences and interactions within their environment. AI chatbots align closely with this theory by offering interactive, student-centered learning opportunities. In this study, these principles were applied by designing chatbot activities that provided personalized feedback, scenario-based learning, and real-time conversations, encouraging students to actively engage with content—approaches which are shown to deepen learning (Piaget, 1971). Prior research, such as Lawasi et al. (2024), has demonstrated that AI-based tools encourage learners to think critically and analytically by

offering multifaceted, scenario-based tasks, echoing Vygotsky's Zone of Proximal Development (ZPD), where chatbots serve as virtual scaffolding. Similarly, Adamopoulou and Moussiades (2020) reviewed multiple chatbot implementations and found that constructivist approaches, like scenario-based dialog and formative feedback, led to meaningful improvements in learner engagement and outcomes.

Self-determination theory (SDT)

Self-Determination Theory posits that motivation is driven by the fulfillment of three basic psychological needs: autonomy, competence, and relatedness (Ryan & Deci, 2000). In this study, SDT informed the evaluation of how chatbots contributed to student motivation and engagement. Chatbots enabled autonomy through self-paced and choice-driven interactions; supported competence via immediate, individualized feedback; and fostered relatedness by simulating conversational exchanges. Research by Chang et al. (2022) supports this approach, showing that chatbot-facilitated environments increased students' self-efficacy and intrinsic motivation by satisfying these psychological needs. Their findings revealed that students felt more in control and confident when interacting with AI chatbots, mirroring the outcomes observed in our study.

Cognitive load theory

Cognitive Load Theory focuses on the limitations of working memory and the importance of minimizing unnecessary cognitive effort (Sweller, 1988). This study used Cognitive Load Theory to assess how chatbots presented information in manageable segments and offered step-by-step guidance, thus reducing extraneous cognitive load. Farrokhnia et al. (2023) reported similar findings, noting that AI chatbots' ability to break down complex tasks and provide targeted support helped learners focus on relevant material and improved retention. This approach aligns with the evidence that reducing cognitive overload enhances both learning outcomes and user satisfaction with digital tools.

Technology acceptance model (TAM)

The Technology Acceptance Model (TAM) explains how users come to accept and use new technologies based on perceived usefulness and ease of use (Davis, 1989). This study incorporated TAM in survey items measuring perceived chatbot usefulness and ease of use, both of which strongly predicted students' actual usage. Gokcearslan et al. (2024) found that students' willingness to use educational chatbots correlated closely with TAM constructs, and that perceived usefulness and user-friendliness were significant drivers of sustained engagement. Okonkwo and Ade-Ibijola (2021) also observed that institutional support and prior experience were key external variables, reinforcing our approach to analyzing attitudes and acceptance.

AI4K12 initiative framework

The AI4K12 Initiative Framework, developed by Touretzky et al. (2019), outlines five core concepts for AI literacy: Perception, Representation and Reasoning, Learning, Natural Interaction, and Societal Impact. This framework was used in our study to ensure that chatbot activities fostered not only language and critical thinking but also broader AI literacy. Zawacki-Richter et al. (2019) emphasizes the importance of integrating AI literacy into educational practice so students can critically interact with AI technologies and understand their ethical and societal dimensions. Our approach reflects these recommendations by promoting understanding of AI's role and impact within the curriculum.

METHODOLOGY

Research Design

This study employs a mixed-methods approach, integrating both quantitative and qualitative methods to provide a comprehensive understanding of the impact of AI chatbots on student learning outcomes at the Diplomatic Academy of Vietnam. This approach allows for a robust analysis by combining numerical data with detailed insights from participants' experiences.

The research was conducted at the Diplomatic Academy of Vietnam (DAV), a leading public institution in Hanoi under the Ministry of Foreign Affairs. The Academy is recognized for its specialized programs in international relations, diplomacy, international law, international economics, communication, and foreign languages. Its primary mission is to train future diplomats, policy analysts, and international affairs professionals, equipping students with strong language proficiency, critical thinking abilities, and cross-cultural communication skills. DAV attracts a diverse cohort of students from across the country, who are selected through competitive entrance exams and often possess high academic achievement and language competency. The academy's learning environment emphasizes academic excellence, professional development, and innovation in teaching methods, making it an ideal setting for investigating the integration of AI-driven educational technologies such as chatbots.

Quantitative methods

The quantitative component of the study focuses on assessing the measurable effects of chatbot usage on language proficiency, critical thinking, and engagement. Data collection involves:

1. Pre- and Post-tests: Administered to evaluate changes in language proficiency and critical thinking skills before and after chatbot integration.

2. Surveys: Designed with Likert-scale items to measure students' perceptions of chatbots in terms of usefulness, ease of use, and engagement.

Qualitative methods

The qualitative component aims to explore participants' perceptions and experiences in greater depth. Data collection includes:

1. Semi-structured Interviews: Conducted with students to gain insights into the benefits and challenges of chatbot integration.

2. Focus Groups: Facilitated discussions with selected participants to understand group dynamics and collective feedback on chatbot usage.

3. Open-ended Survey Questions: Providing respondents the opportunity to share detailed, narrative responses.

Rationale for mixed-methods approach

The mixed-methods design is chosen to address the study's research questions comprehensively:

1. Quantitative data offers objective, statistically valid insights into learning outcomes and user perceptions.

2. Qualitative data provides contextual understanding, uncovering nuanced experiences and opinions that numbers alone cannot capture.

3. The integration of these methods ensures the findings are both generalizable and richly descriptive, enabling actionable recommendations for the effective implementation of AI chatbots in educational settings.

This dual approach ensures a balanced and in-depth exploration of the research objectives, providing evidence-based conclusions that are both statistically sound and contextually relevant.

Participants

Student participants

The study includes 520 undergraduate students from the Diplomatic Academy of Vietnam. These students were enrolled in courses where AI chatbots were integrated as part of the learning process. Participants represent various fields of study, including international relations, international economics, communication, and linguistics. Their English proficiency levels range from intermediate to advanced (CEFR B1-C1).

Sampling method

A purposive sampling method was employed to ensure that participants

were directly involved in courses utilizing AI chatbots. This approach enabled the study to focus on individuals most affected by the technology, providing targeted insights into its effectiveness and challenges.

Criteria for selection:

Participants were selected based on the following criteria:

- Enrollment in undergraduate courses at the Diplomatic Academy of Vietnam where AI chatbots were actively integrated into the curriculum.
- Representation across multiple academic disciplines (including international relations, international economics, communication, and linguistics).
- English proficiency at an intermediate to advanced level (CEFR B1–C1), to ensure effective engagement with chatbot-based language tasks.
- Willingness to provide informed consent for participation in the research.

Participant Selection Process for Each Method:

- Survey:

  All students (N = 520) enrolled in courses utilizing AI chatbots during the study period were invited to participate in the structured survey. This broad inclusion ensured comprehensive quantitative data on user perceptions, frequency of chatbot use, and perceived effectiveness.

- Pre- and post-tests:

  The same 520 students who participated in the survey also completed pre- and post-tests, which assessed changes in language proficiency and critical thinking skills. These assessments were administered at the start and end of the academic semester to all eligible students.

- Interviews:

  From the survey respondents, approximately 30 participants were purposively selected for semi-structured interviews. Selection aimed for diversity in academic year, major, and reported frequency of chatbot use, ensuring a range of perspectives and experiences were captured.

- Focus groups:

  Three focus groups were conducted, each consisting of 6–8 participants (total n ≈ 20). Focus group participants were chosen from survey respondents who expressed willingness to engage in group discussions, with an emphasis on achieving a mix of academic backgrounds and chatbot engagement levels.

- Open-ended survey questions:

  All 520 students who completed the structured survey were also invited to respond to open-ended survey questions, allowing for detailed narrative responses about their experiences with AI chatbots.

Instrument	Number of Participants	Selection Criteria/Process
Survey	520	All students in chatbot-integrated courses
Pre- and Post-Tests	520	Same as survey participants
Interviews	30	Purposive selection for diversity (major, year, chatbot use)
Focus Groups	18–24 (3 x 6–8)	Volunteers from survey, mixed academic backgrounds
Open-ended Survey Questions	520	Same as survey participants

Demographic details

Students are chosen from first-year to final-year, with a balanced gender distribution. They represent diverse academic disciplines and educational backgrounds. This comprehensive participant pool ensures that the study captures a wide range of experiences and perceptions, offering a nuanced understanding of AI chatbot integration in an academic setting.

Data Collection

Quantitative data collection

- Surveys

The structured survey administered in this study was developed based on an extensive review of the literature on AI chatbot integration in education (Adamopoulou & Moussiades, 2020; Gokcearslan et al., 2024; Okonkwo & Ade-Ibijola, 2021). Items were adapted from previously validated instruments that measured technology acceptance, language learning outcomes, critical thinking, and student engagement. The Likert-scale questions were designed to capture participants' perceived improvements, frequency of use, and attitudes toward AI chatbots in comparison with traditional methods.

To ensure content validity, the initial survey draft was reviewed by a panel of subject matter experts, including faculty members specializing in language pedagogy, educational technology, and assessment at the Diplomatic Academy of Vietnam. Feedback from these experts was incorporated to refine question wording, response scales, and coverage of relevant domains. A pilot survey was then conducted with a small group of students (n=20) not included in the main sample, to assess clarity, reliability, and completion time. Minor adjustments were made based on pilot feedback, resulting in a final survey instrument with strong face and content validity.

- Pre- and Post-tests

The pre- and post-tests were designed to provide objective measures of students' language proficiency and critical thinking skills. Test items were adapted from standardized English proficiency assessments (aligned with CEFR B1–C1 levels) and validated critical thinking rubrics commonly used in higher education. The language component included writing prompts, vocabulary usage, and listening/speaking simulation tasks relevant to the diplomatic context, while the critical thinking section presented scenario-based and open-ended problem-solving tasks.

To ensure the validity and reliability of these assessments:

- Test content was reviewed and validated by experienced instructors in the Academy's English and international relations departments.
- Test blueprints were aligned with course learning outcomes to ensure construct validity.
- A pilot administration of the tests was conducted with a subset of students to evaluate item difficulty, clarity, and time requirements.
- Statistical analysis (such as item analysis and calculation of Cronbach's alpha for internal consistency) was performed on pilot data, and revisions were made to enhance reliability.

Qualitative data collection

- Open-ended survey questions

The online surveys included open-ended questions, allowing students to elaborate on their experiences and perceptions regarding AI chatbot usage. These questions encouraged detailed responses on both the benefits and challenges of chatbot integration. Examples of prompts included inquiries such as how AI chatbots enhanced their critical thinking abilities and which chatbot features

they found most engaging or helpful. The responses were analyzed thematically to identify patterns in user experiences, providing valuable insights into students' perceptions of chatbot effectiveness and engagement.

- Semi-structured interviews

In addition to survey responses, semi-structured interviews were conducted with approximately 30 participants to explore their experiences in greater depth. These interviews focused on topics such as the benefits and limitations of chatbot use, specific scenarios where chatbots contributed to learning improvements, and suggestions for optimizing their integration into academic settings. The semi-structured format ensured consistency in key discussion areas while allowing students to elaborate on their perspectives, offering rich qualitative data on chatbot effectiveness in different learning contexts.

- Focus groups

Furthermore, three focus groups, each consisting of six to eight participants, were organized to facilitate discussions on collaborative learning experiences involving AI chatbots. The discussions examined how chatbots influenced peer interactions, supported group projects, and enhanced collaborative learning tasks. By analyzing shared experiences and group dynamics, focus group discussions provided additional context on how chatbots shaped student collaboration and engagement.

- Ethical considerations

This study adhered to strict ethical guidelines to ensure the rights and well-being of all participants. Prior to participation, all students were fully briefed on the study's objectives, methodology, and potential implications. Informed consent was obtained from each participant, ensuring that they were aware of their rights and the voluntary nature of their involvement. To maintain confidentiality,

all data collected was anonymized, preventing the identification of individual respondents and safeguarding their privacy. Additionally, participants retained the right to withdraw from the study at any stage without facing any consequences or penalties.

Data Analysis

Quantitative data analysis

- Descriptive statistics

Descriptive statistics were used to summarize the data collected through surveys and pre- and post-tests, providing an overview of chatbot usage patterns and learning outcomes. Key metrics included means, standard deviations, and frequency distributions, which helped illustrate variations in student engagement and perceived improvements in language proficiency and critical thinking.

- Inferential statistics

Inferential statistical methods were applied to determine the significance of observed changes. Paired t-tests were conducted to analyze differences between pre- and post-test scores, assessing whether chatbot usage had a statistically significant effect on students' language proficiency and critical thinking skills. Furthermore, correlation analyses were performed to examine potential relationships between chatbot usage frequency and perceived engagement levels.

Qualitative data analysis

- Thematic analysis

A thematic analysis was conducted to examine data from semi-structured interviews, focus groups, and open-ended survey responses. Key themes identified included the perceived benefits of AI chatbots, such as increased engagement and accessibility, as well as challenges like technical issues and over-reliance on

technology. Additionally, participants provided suggestions for improvement, including enhanced personalization and better integration into learning environments.

- Coding process

To ensure a systematic approach, the coding process involved analyzing transcripts and narrative responses using NVivo software. Recurring patterns and unique insights were identified, with codes grouped into broader themes aligned with the study's research objectives. This structured analysis provided a comprehensive understanding of students' experiences and perceptions of AI chatbot integration in education.

Triangulation

To enhance the validity and reliability of the findings, a triangulation approach was employed by cross-verifying quantitative and qualitative data. Improvements in pre- and post-test scores were analyzed alongside students' self-reported perceptions of skill development to ensure consistency between objective and subjective measures. Additionally, qualitative insights from interviews and focus groups were used to contextualize survey results, providing a deeper understanding of students' experiences with AI chatbots.

To ensure the trustworthiness of the qualitative findings, several measures of validity and reliability were incorporated throughout the research process:

- Credibility: The credibility of the thematic analysis was supported by member checking, where selected participants were invited to review and verify interpretations of their responses for accuracy. Researcher triangulation was also employed, with multiple members of the research team independently coding a subset of transcripts to confirm consistency in theme identification and interpretation.

- Dependability: An audit trail was maintained throughout the data analysis process, documenting coding decisions, theme development, and any changes to the analytic framework. This transparency allows the analysis to be traced and replicated by other researchers.
- Confirmability: NVivo software was used to systematically organize and code the qualitative data, minimizing researcher bias and ensuring that findings were grounded in the data rather than personal interpretation. Reflexivity was practiced, with researchers keeping analytic memos to record and reflect on potential influences on the interpretation process.
- Transferability: Thick descriptions of the research setting, participant demographics, and context were provided, enabling readers to determine the applicability of the findings to other educational settings or contexts.

By implementing these procedures, the study enhanced the rigor, trustworthiness, and credibility of its qualitative analysis, providing robust and reliable insights into students' experiences with AI chatbot integration.

RESULTS AND DISCUSSION

Impact on Language Proficiency

Quantitative findings

Data from pre- and post-tests revealed significant improvements in students' language proficiency after the integration of AI chatbots into their learning. The mean pre-test score was 65.32%, which increased to 80.47% in the post-test, reflecting a 15.15% improvement. These findings were confirmed through paired t-tests (t-statistic = 45.23, $p < 0.001$), indicating statistically significant gains across all skill categories.

Table 1: Pre- and Post-test Score Comparison

Language Skill	Pre-test (%)	Post-test (%)	Improvement (%)
Listening & Speaking	65.32	80.47	15.15
Writing	60.50	75.40	14.90
Overall	62.5	77.5	15.15

The improvement in listening and speaking reflects the effectiveness of chatbots in simulating conversational practice. Similarly, the progress in writing can be attributed to chatbots' ability to provide instant grammar corrections and tailored feedback.

Survey findings

Findings from surveys conducted with 520 students further support the positive impact of AI chatbots on language development. A substantial 76.2% of students reported an overall improvement in language proficiency, reflecting gains across multiple language skills. Notably, 55% of respondents highlighted progress in listening and speaking, particularly in pronunciation and fluency. Additionally, 67% noted enhancements in writing, with improvements in vocabulary range and grammatical accuracy.

Table 2: Survey-reported Improvement in Language Skills

Language Skill	Percentage of Students Reporting Improvement (%)
Listening & Speaking	55
Writing	67
Overall	76.2

Students consistently noted that the chatbot exercises provided personalized feedback, allowing them to target areas of difficulty effectively. These results align with the statistical gains from pre- and post-tests, further validating the perceived benefits of chatbot integration.

Qualitative insights

Themes emerging from interviews and open-ended survey responses provided deeper insights into the mechanisms driving student improvements with AI chatbots. A recurring theme was increased confidence, as many students viewed chatbots as judgment-free platforms for practicing English, reducing the anxiety typically associated with speaking in front of peers. One student noted, "ChatGPT provided a safe space for practicing conversations, making me more confident in speaking English."

Another key insight was the benefit of real-time feedback, where the immediacy of chatbot corrections enabled students to identify and rectify errors instantly, accelerating their progress. As one student shared, "The chatbot helped me learn better alternatives for sentence construction and improve faster."

Additionally, students highlighted the advantage of tailored learning experiences, as chatbots' adaptive capabilities allowed them to focus on specific weaknesses, such as grammar or advanced vocabulary. One participant remarked, "I could practice at my own pace and concentrate on difficult areas." These qualitative insights contextualized the quantitative findings, illustrating how chatbot interactions contributed to students' skill development in language learning.

Synthesis of findings

The convergence of quantitative and qualitative findings highlights the significant impact of AI chatbots on language proficiency. Quantitative evidence from pre- and post-test scores provides objective validation of students' skill enhancement, demonstrating measurable improvements in language proficiency and critical thinking. Complementing this, student perceptions gathered through surveys and interviews emphasize the practical and emotional benefits of chatbot use, including increased confidence and reduced anxiety in language practice.

These findings also align with key theoretical frameworks. Constructivist Learning Theory supports the role of AI chatbots in facilitating active, interactive learning, while Self-Determination Theory underscores how chatbots promote autonomy and competence, motivating students to engage in self-directed learning. By integrating both objective performance data and subjective experiences, this study provides a comprehensive understanding of the pedagogical value of AI chatbots in diplomatic education.

Impact on Critical Thinking

Quantitative findings

The pre- and post-test results revealed measurable improvements in students' critical thinking skills after chatbot integration. The mean critical thinking score increased from 58.40% in the pre-test to 73.55% in the post-test, reflecting a 15.15% improvement. Paired t-tests confirmed the statistical significance of this improvement (t-statistic = 39.87, $p < 0.001$), emphasizing the effectiveness of chatbots in enhancing analytical and evaluative abilities.

Table 3: Pre- and Post-test Score Comparison for Critical Thinking

Metric	Pre-test (%)	Post-test (%)	Improvement (%)
Critical Thinking	58.40	73.55	15.15

The substantial improvement suggests that chatbot exercises, particularly scenario-based and open-ended tasks, were instrumental in building students' reasoning and decision-making capabilities.

Survey findings

Survey responses further validate the role of AI chatbots in fostering critical thinking skills. A significant 71% of students reported an improvement in their ability to analyze and evaluate information after engaging with chatbot-based activities. Additionally, students identified specific chatbot features, such as scenario simulations and reflective

prompts, as particularly effective in enhancing their reasoning and problem-solving abilities.

Table 4: Survey-reported Improvement in Critical Thinking

Feature	Percentage of Students Reporting Effectiveness (%)
Scenario-based Tasks	68
Reflective Prompts	62
Open-ended Problem Solving	59

Students found scenario-based tasks particularly engaging, as they required evaluating multiple perspectives and formulating reasoned conclusions. Reflective prompts encouraged metacognitive thinking, helping learners to critically assess their own reasoning processes.

Qualitative insights

Student interviews and open-ended survey responses provided deeper insights into how AI chatbots fostered critical thinking skills. A key theme was scenario-based learning, where chatbots presented real-world challenges that required students to analyze data, evaluate alternatives, and propose solutions. As one student noted, "The chatbot's tasks made me think about different outcomes and consider the best decisions."

Another significant finding was the role of reflective prompts in encouraging deeper reasoning. Chatbots prompted students to justify their responses and reconsider their thought processes, reinforcing higher-order thinking skills. One participant explained, "After answering, the chatbot asked why I chose that solution, which made me think deeper."

Additionally, the interactive nature of chatbot discussions stimulated analytical thinking by challenging students to explore multiple perspectives. The dynamic exchange of ideas encouraged students to refine their arguments and consider alternative viewpoints. As one student described, "The chatbot challenged my opinions and encouraged me to explain my reasoning."

Synthesis of findings

The integration of quantitative and qualitative data underscores the significant role AI chatbots play in fostering critical thinking skills. Data validation is evident as statistical improvements in test scores align with survey responses, confirming students' perceptions of enhanced analytical abilities. The practical application of chatbots is further reinforced through scenario-based tasks and reflective prompts, which simulate real-life decision-making processes and equip students with skills transferable to academic and professional contexts. These findings also demonstrate theoretical alignment with Constructivist Learning Theory, which emphasizes active, problem-centered learning, and Cognitive Load Theory, which highlights the importance of scaffolding complex tasks to optimize cognitive processing. Together, these insights validate the effectiveness of AI chatbots as a tool for developing critical thinking in educational settings.

Impact on Engagement

Quantitative findings

Engagement levels were assessed through survey responses and chatbot usage data, revealing a strong positive impact on student participation. A significant 81.5% of students reported increased engagement with their studies after using AI chatbots. This heightened engagement was reflected in chatbot interaction patterns, with 45% of students using the chatbot daily, primarily to practice conversational skills. Additionally, 32% engaged 2–3 times per week, focusing on grammar and vocabulary exercises.

Table 5: Frequency of Chatbot Usage

Frequency of Use	Percentage of Students (%)
Daily	45
2-3 times per week	32
Once per week or less	23

The surveys also highlighted that 78% of students found the chatbot's interactive features, such as instant feedback and gamified tasks, motivating. These features were credited with making learning enjoyable and reducing the monotony often associated with traditional study methods.

Survey-reported engagement drivers

Survey responses identified key factors contributing to increased student engagement with AI chatbots. Instant feedback emerged as the most engaging feature, with 72% of students recognizing its role in providing immediate corrections and guidance. Personalized learning, highlighted by 69% of respondents, was valued for its tailored tasks and adaptive feedback that addressed individual learning needs. Additionally, gamified features were cited by 63% of students as a motivating element, making learning more interactive and enjoyable.

Table 6: Engagement-driving Features

Feature	Percentage of Students Reporting Effectiveness (%)
Instant feedback	72
Personalized learning	69
Gamified tasks	63

Qualitative findings

Interviews and open-ended survey responses highlighted that students found chatbot interactions engaging due to their interactive and adaptive nature. One key theme was interactive and adaptive learning, where students appreciated how chatbots adjusted to their progress by offering challenges that matched their skill levels. As one student explained, "The chatbots kept me interested because they adjusted the difficulty as I improved."

Another significant factor was the creation of an enjoyable learning environment through gamified elements such as scoring systems and progression badges, which made learning feel less like a chore and more like an engaging activity. One student remarked, "It didn't feel like studying. The games and challenges motivated me to keep practicing."

Additionally, students emphasized the 24/7 accessibility of chatbots as a major engagement driver. The ability to practice at any time allowed them to integrate language learning into their daily routines more easily. One participant noted, "Having the chatbots available whenever I needed them meant I could practice more often without worrying about a fixed schedule."

Synthesis of findings

The integration of quantitative and qualitative data provides a comprehensive picture of how chatbots enhance student engagement. Frequent usage rates reflect students' willingness to engage, driven by interactive features and accessibility. Student feedback from surveys and qualitative responses highlights the role of instant feedback and gamification in sustaining motivation. The findings align with the Technology Acceptance Model (TAM), as students' perceptions of usefulness and ease of use influenced engagement. The self-paced nature of chatbot learning also supports Self-Determination Theory (SDT) by fulfilling students' needs for autonomy and competence.

Perceptions of Students

Benefits of chatbots

Survey results and qualitative feedback highlighted several key benefits of chatbot integration in learning. Enhanced learning experience was a major advantage, with 78% of students agreeing that chatbots made lessons more interactive and accessible, providing opportunities for consistent, real-time practice. Another significant benefit was instant feedback, as 72% of respondents valued the immediate corrections and

suggestions, which helped refine grammar, vocabulary, and sentence structure. Additionally, personalized learning pathways were highly appreciated, with 69% of students recognizing the adaptability of chatbots in tailoring exercises to meet their individual learning needs.

Table 7: Perceived Benefits of AI Chatbots

Perceived Benefit	Percentage of Students (%)
Enhanced Learning Experience	78
Instant Feedback	72
Personalized Learning	69

Qualitative responses reinforced these findings, highlighting the adaptability and accessibility of chatbots in learning. One student noted, "The chatbot adjusted to my weaknesses, giving me exercises that helped me improve faster." Another remarked, "It felt like having a tutor available whenever I needed help."

Challenges of chatbots

Despite positive feedback, students reported several challenges. Limited feedback accuracy was noted by 35% of respondents, who found chatbot responses repetitive or ineffective for complex queries. Technical issues were highlighted by 30% of students, citing reliance on a stable internet connection and occasional performance glitches.

Table 8: Challenges Reported by Students

Perceived Challenge	Percentage of Students (%)
Limited Feedback Accuracy	35
Technical Issues	30

Some students noted that repetitive feedback diminished the learning experience. One student remarked, "Sometimes the chatbot kept repeating the same advice without helping me learn more effectively." This highlights a limitation in chatbot responsiveness and adaptability.

Practical insights

Students provided practical recommendations for improving chatbot integration. Enhancing feedback quality was a recurring suggestion, with calls for more varied and contextually accurate responses. Integration into the curriculum was also recommended, with students advocating for a balanced approach that combines chatbot exercises with traditional lessons. Additionally, technical improvements were suggested, particularly in ensuring system stability and reducing reliance on high-speed internet.

Comparison with Existing Studies

Alignment with prior research

The findings of this study align with existing research on AI chatbots in education, particularly in language proficiency, critical thinking, and engagement.

Language Proficiency: The observed improvement in language proficiency supports the findings of Hill et al. (2015), who demonstrated that chatbots enhance vocabulary acquisition and conversational skills through interactive dialogues. Similarly, Farrokhnia et al. (2023) highlighted that adaptive feedback from AI tools improves grammar and writing fluency. The immediate feedback feature, identified as a key driver of improvement by 72% of students, mirrors the conclusions of Adamopoulou and Moussiades (2020), who emphasized chatbots' role in creating personalized, responsive learning environments.

Critical Thinking: The evidence that scenario-based tasks and reflective prompts foster critical thinking aligns with Lim and Makany's (2023) research, which demonstrated that chatbots engage students in higher-order thinking by presenting complex, reflective tasks. Similarly, Chen et al. (2024) found that chatbots promote metacognitive thinking by prompting learners to evaluate their reasoning, consistent with this study's qualitative findings.

Engagement: The high engagement levels reported in this study, with 81.5% of students indicating increased motivation, align with Fryer et al. (2019), who found that gamified chatbot interactions enhance student motivation and participation. The results also echo Kumar (2021), who noted that chatbots foster learner engagement through personalized and adaptive interactions.

Differences from prior research

This study presents several distinctions from existing research, particularly in domain-specific applications, over-reliance on AI, and technical limitations.

Domain-Specific Applications: While studies such as Adamopoulou and Moussiades (2020) focus on general education contexts, this study uniquely examines AI chatbot integration in diplomacy education. The findings indicate that chatbots require further customization to address domain-specific skills, such as formal writing and professional communication, a gap less emphasized in prior research.

Over-Reliance on AI: Unlike Lim and Makany's (2023) research, which primarily highlighted chatbot benefits, this study revealed concerns about over-reliance on AI tools. Some students expressed apprehension that excessive dependence on chatbots might limit their ability to explore other learning methods, underscoring the need for a balanced integration of AI with traditional approaches.

Technical Limitations: While Hill et al. (2015) placed less emphasis on technical challenges, this study identified inconsistent internet connectivity and occasional chatbot errors as notable barriers. These findings suggest that infrastructure and technology readiness play a critical role in the successful implementation.

Theoretical Discussion of Findings

The results of this study not only align with, but also extend, the current body of research on AI chatbots in education through the application of several key educational theories presented in the literature review.

Constructivist learning theory

The improvements in language proficiency and critical thinking observed in this study can be interpreted through the lens of constructivist learning (Piaget, 1971). By engaging students in scenario-based tasks, real-time dialogue, and personalized feedback, chatbots enabled learners to construct knowledge actively rather than passively receive it. The study's qualitative data, students reporting increased confidence and the ability to practice conversational skills without judgment, mirrors Vygotsky's notion of the Zone of Proximal Development, where chatbots acted as virtual scaffolds, providing support just beyond the learners' current abilities (Lawasi et al., 2024). This supports the conclusion that constructivist-aligned chatbot activities facilitate deeper, more meaningful learning experiences.

Self-determination theory (SDT)

The marked increase in student engagement and motivation can be attributed to the satisfaction of the SDT's three basic psychological needs: autonomy, competence, and relatedness (Ryan & Deci, 2000). Chatbots promoted autonomy by enabling self-paced learning and choice in task selection, competence through immediate and personalized feedback, and relatedness by simulating supportive conversational partners. These findings echo Chang et al. (2022), who reported that chatbot environments fostering these needs enhance intrinsic motivation, and are further supported by students' self-reported growth in confidence and engagement.

Cognitive load theory

Cognitive Load Theory (Sweller, 1988) offers a useful framework for understanding how chatbots contributed to improved comprehension and retention. By breaking complex information into manageable segments and providing step-by-step guidance, chatbots helped reduce extraneous cognitive load, allowing students to focus their mental resources on essential learning tasks. The high rates of student-reported

ease of use and clarity, along with improved post-test scores, suggest that chatbots can be a valuable tool for optimizing instructional design and minimizing cognitive barriers to learning.

Technology acceptance model (TAM)

The widespread adoption and frequent usage of chatbots reported in this study align with the Technology Acceptance Model (Davis, 1989), which emphasizes perceived usefulness and ease of use as critical factors for technology integration. Survey data showing high levels of satisfaction and willingness to continue using chatbots confirm TAM's relevance. Notably, the study's identification of external barriers — such as technical limitations and the need for institutional support — also aligns with TAM's assertion that external variables influence user acceptance and sustained engagement.

AI4K12 initiative framework

Finally, the integration of the AI4K12 Initiative Framework (Touretzky et al., 2019) in chatbot activities ensured that students were not only developing linguistic and cognitive skills, but also building AI literacy. The findings indicate that chatbots can serve as a bridge to help students understand core AI concepts, such as natural language processing and ethical considerations in AI use — an outcome increasingly vital for future diplomats and international professionals (Zawacki-Richter et al., 2019). The study's emphasis on domain-specific applications and the need for customization further reinforces the importance of contextualizing AI tools within the broader framework of AI literacy and societal impact.

CONCLUSION

Summary of Key Findings

Research Question 1: Impact of AI Chatbots on Language Proficiency, Critical Thinking, and Engagement

The findings of this study indicate a significant impact of AI chatbots on students' language proficiency, critical thinking, and engagement. In terms of language proficiency, students demonstrated a 15.15% improvement in pre- and post-test scores, with notable gains in listening, speaking, and writing skills. The primary factors contributing to this progress included real-time feedback, personalized exercises, and simulated conversational practice, which collectively enhanced grammar accuracy, vocabulary acquisition, and overall fluency. These results suggest that AI chatbots provide an effective and interactive platform for language learning by facilitating consistent practice and immediate correction.

Regarding critical thinking skills, AI chatbots played a crucial role in fostering analytical reasoning through scenario-based tasks and reflective prompts. These interactive elements encouraged students to analyze information, evaluate different perspectives, and justify their reasoning, thereby strengthening their cognitive abilities. The improvement in critical thinking was further supported by a statistically significant increase in test scores, coupled with student feedback emphasizing the value of reflective activities. This suggests that chatbots can serve as effective tools for promoting higher-order thinking skills when properly integrated into academic settings.

The study also found a strong positive correlation between AI chatbot use and student engagement, with 81.5% of participants reporting increased motivation and participation in learning activities. The interactive features of chatbots, including instant feedback, gamified tasks, and 24/7 accessibility, contributed significantly to this heightened engagement. These elements transformed the learning process into an enjoyable and interactive experience, which, in turn, encouraged frequent participation and consistent practice. The results confirm that AI chatbots have the potential to sustain student engagement through dynamic and personalized learning experiences.

Research Question 2: Benefits and Challenges of AI Chatbot Integration

The study identified several key benefits of integrating AI chatbots into academic settings. One of the most frequently reported advantages was accessibility and convenience, as students valued the ability to practice at any time, enabling consistent learning without time constraints. Another significant benefit was personalized learning, with chatbots tailoring exercises to address individual weaknesses, making the learning process more targeted and efficient. Additionally, students noted that chatbot interactions contributed to confidence building, particularly in speaking and writing tasks, by creating a non-judgmental environment that reduced anxiety. Furthermore, the inclusion of gamification, progress tracking, and interactive elements in chatbot design made learning more engaging and motivating, encouraging students to maintain a steady learning routine.

Despite these benefits, students also highlighted certain challenges associated with chatbot integration. One of the primary concerns was technical limitations, as some students encountered occasional chatbot errors and internet dependency issues, which disrupted their learning experience. Another notable challenge was repetitive feedback, where students found that chatbot responses sometimes lacked depth and failed to address complex queries effectively. Additionally, a minority of students expressed concerns regarding over-reliance on AI, fearing that excessive dependence on chatbots might lead to a neglect of other learning resources and interpersonal interactions. These findings suggest that while AI chatbots offer numerous advantages, their effective integration into academic environments requires continuous improvement in response accuracy, technical stability, and balance with traditional learning methods.

Implications for Practice

The findings of this study highlight the significant potential of AI chatbots to enhance language proficiency, critical thinking, and engagement

among students. However, to maximize their effectiveness and address identified challenges, several recommendations for integrating chatbots into academic settings can be made.

First, improving the quality of chatbot feedback is essential. While students appreciated the immediate corrections provided by chatbots, many reported that feedback on advanced tasks was repetitive or lacked depth. Enhancing the contextual accuracy and variety of chatbot responses can better support students in tackling complex learning tasks. This could be achieved by leveraging advanced natural language processing models and incorporating task-specific prompts that challenge higher-order thinking skills.

Addressing technical limitations is another crucial step. Students frequently cited performance issues, including dependence on stable internet connections and occasional glitches, as barriers to seamless learning experiences. To overcome this, educational institutions and developers should prioritize creating offline-compatible chatbot applications and conducting regular updates to ensure reliability. A more robust technical infrastructure will reduce frustration and make chatbot-based learning more accessible.

The study also underscores the importance of balancing chatbot usage with traditional teaching methods. Over-reliance on chatbots may limit opportunities for interpersonal interaction and collaborative learning. Therefore, chatbots should be integrated into a blended learning framework, where they complement instructor-led activities and peer discussions. For example, chatbots can serve as tools for independent practice, while instructors focus on providing human insights and facilitating group tasks that promote social interaction.

Personalizing chatbot interactions further enhances their appeal and effectiveness. Students valued the adaptability of chatbots in tailoring exercises to their individual needs, and expanding this capability can strengthen engagement and outcomes. By leveraging data analytics to track student performance, chatbots can dynamically adjust task

difficulty and align with learners' specific goals. Allowing students to choose task types and levels of complexity would also empower them to take ownership of their learning journey.

Gamification emerged as a key driver of engagement in this study. Features such as progress tracking, rewards, and interactive challenges motivated students to practice regularly and stay committed to their learning goals. Strengthening these gamified elements by introducing leaderboards, milestone rewards, and customizable features can further sustain long-term interest and motivation.

Customizing chatbots for domain-specific applications is another area for improvement. This study highlights the need for chatbots tailored to specialized fields such as diplomacy, where learners require unique skill sets like formal writing and negotiation. Developers should collaborate with subject-matter experts to create content that reflects real-world scenarios in specialized domains, ensuring that chatbot tasks are both relevant and practical.

Finally, providing training for educators is vital for the successful integration of chatbots into academic contexts. Teachers play a crucial role in guiding students to use chatbots effectively and purposefully. Workshops and training sessions can equip educators with the knowledge and skills to incorporate chatbots into their teaching practices. Additionally, developing manuals and best-practice guidelines can help instructors integrate chatbot activities seamlessly into their curriculum.

Limitations of the Study

While this study provides valuable insights into the impact of AI chatbots on student learning outcomes at the Diplomatic Academy of Vietnam, several limitations must be acknowledged. These constraints highlight areas that could be addressed in future research to ensure a more comprehensive understanding of the topic.

One key limitation lies in the participant pool. The study focused exclusively on 520 undergraduate students from a single institution, which may limit the generalizability of the findings. While the sample was diverse in terms of academic disciplines, including international relations, economics, and communication, the findings may not fully reflect the experiences of students in other educational contexts, such as vocational or technical training programs. Expanding the study to include participants from multiple institutions and varied educational settings could provide a broader perspective on the effectiveness of AI chatbots.

The study's reliance on self-reported data is another constraint. While surveys and interviews provided valuable insights into students' perceptions, self-reported data can be influenced by social desirability bias, where participants may overstate the benefits or downplay challenges of chatbot usage. Incorporating objective measures, such as analyzing interaction logs from chatbots or conducting longitudinal studies, could provide a more balanced understanding of their impact.

Time constraints also posed a limitation. The study was conducted over two academic semesters, which may not have been sufficient to capture the long-term effects of chatbot usage on learning outcomes. For example, while immediate improvements in language proficiency and critical thinking were observed, the sustainability of these gains over time remains uncertain. Future research could explore the longitudinal effects of chatbot integration, examining how continued usage impacts skills retention and transfer to real-world applications.

Additionally, the study did not fully explore how individual differences, such as technological proficiency or learning preferences, influenced the effectiveness of chatbots. While most students reported positive experiences, it is possible that some learners found chatbots less intuitive or engaging due to varying comfort levels with technology. Investigating these individual factors could help in tailoring chatbot features to accommodate diverse learner profiles.

Another limitation involves the specific features of the chatbots used. The study primarily relied on general-purpose AI tools like ChatGPT, which may not have been optimized for educational purposes. The absence of domain-specific content or advanced pedagogical features may have limited the potential benefits of the chatbots. Future studies could examine the impact of customized educational chatbots designed to address the unique needs of learners in specialized fields such as diplomacy.

Finally, the study's geographic focus on Vietnam provides valuable insights into the local educational context but may limit the applicability of findings to other cultural or linguistic settings. Cultural attitudes toward technology and learning styles can significantly influence the acceptance and effectiveness of AI tools. Expanding research to include cross-cultural comparisons could provide a more comprehensive understanding of how chatbots perform in diverse educational environments.

Suggestions for Future Research

Future studies should explore the long-term impacts of AI chatbot integration on student learning outcomes. While this study highlighted immediate improvements in language proficiency, critical thinking, and engagement, it remains unclear whether these gains are sustainable over time. Longitudinal research could investigate how chatbot-assisted learning translates into real-world applications, such as professional communication or skill retention.

Another key area for exploration is cross-cultural research. Different cultural and linguistic contexts may influence how students interact with and benefit from chatbots. Investigating these differences could help identify universal features of effective chatbot integration while uncovering the need for cultural adaptations to optimize learning in diverse settings.

Customizing chatbots for specialized domains is also essential. This study primarily used general-purpose chatbots, which may not fully address the unique requirements of fields like diplomacy, healthcare, or law. Developing domain-specific chatbots with tailored content and scenarios can better support students in acquiring contextually relevant skills.

Additionally, future research should compare the effectiveness of chatbot-based learning with traditional instructor-led methods. Such studies would provide insights into the optimal balance between AI and human interaction in education, ensuring that chatbots complement rather than replace traditional approaches.

Lastly, ethical and psychological implications warrant further attention. Over-reliance on chatbots, potential impacts on student well-being, and concerns about data privacy are critical areas for investigation. Addressing these issues will help ensure that chatbots are integrated responsibly and support a balanced, ethical learning experience.

REFERENCES

Adamopoulou, E., & Moussiades, L. (2020). *An overview of chatbot technology: Design, applications, and challenges.* In I. Maglogiannis, L. Iliadis, J. Macintyre, M. Avlonitis, & A. Papaleonidas (Eds.), *Artificial Intelligence Applications and Innovations* (IFIP AICT 712, pp. 373–383). Springer. https://doi.org/10.1007/978-3-030-49186-4_31

Ait Baha, T., El Hajji, M., Es-saady, Y., & Fadili, H. (2022). Towards highly adaptive Edu-Chatbot. *Procedia Computer Science, 198*, 397–403. https://doi.org/10.1016/j.procs.2021.12.260

Al-Zahrani, A. M. (2023). *The impact of generative AI tools on researchers and research: Implications for academia in higher education.* Innovations in Education and Teaching International, 1-15. https://doi.org/10.1080/14703297.2023.2271445

Cao, C. C., Ding, Z., Lin, J., & Hopfgartner, F. (2023). *AI chatbots as multirole pedagogical agents: Transforming engagement in CS education* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2308.03992

Chang, C.-Y., Hwang, G.-J., & Gau, M.-L. (2022). Promoting students' learning achievement and self-efficacy: A mobile chatbot approach for nursing training. *British Journal of Educational Technology, 53*(1), 171–188. https://doi.org/10.1111/bjet.13158

Chen, S., Cheng, H., & Huang, Y. (2024). Emotion recognition in self-regulated learning: Advancing metacognition through AI-assisted reflections. In D. Kourkoulou, A. O. Tzirides, B. Cope, & M. Kalantzis (Eds.), *Trust and inclusion in AI-mediated education* (pp. 188–191). Springer. https://doi.org/10.1007/978-3-031-64487-0_9

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–340. https://doi.org/10.2307/249008

Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2023). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International, 61*(3), 460–474. https://doi.org/10.1080/14703297.2023.2195846

Fryer, L. K., Nakao, K., & Thompson, A. (2019). Chatbot learning partners: Connecting learning experiences, interest, and competence. *Computers in Human Behavior, 93*, 279–289. https://doi.org/10.1016/j.chb.2018.12.023

Gokcearslan, S., Tosun, C., & Erdemir, Z.G. (2024). Benefits, challenges, and methods of Artificial Intelligence (AI) chatbots in education: A systematic literature review. *International Journal of Technology in Education (IJTE), 7*(1), 19-39. https://doi.org/10.46328/ijte.600

Hill, J., Ford, W. R., & Farreras, I. G. (2015). Real conversations with artificial intelligence: A comparison between human–human online conversations and human–chatbot conversations. *Computers in Human Behavior, 49*, 245–250. https://doi.org/10.1016/j.chb.2015.02.026

Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign. ISBN13: 978-1794293700.

Kumar, J. A. (2021). Educational Chatbots for Project-Based Learning: Investigating Learning Outcomes for a Team-Based Design Course. *International Journal of Educational Technology in Higher Education, 18,* Article No. 65. https://doi.org/10.1186/s41239-021-00302-w

Lawasi, M. C., Rohman, V. A., & Shoreamanis, M. (2024). The use of AI in improving student's critical thinking skills. *Proceedings Series on Social Sciences & Humanities, 18*, 366–370. https://doi.org/10.30595/pssh.v18i.1279

Lim, S. S., & Makany, T. (2023). Deploying chatbots to build students' critical thinking skills: Leveraging generative AI effectively and purposefully in higher education. In M. A. Peters & R. Heraud (Eds.), *Encyclopedia of educational innovation*. Springer.

Okonkwo, C. W., & AdeIbijola, A. (2021). Chatbots applications in education: A systematic review. *Computers and Education: Artificial Intelligence, 2*, Article 100033. https://doi.org/10.1016/j.caeai.2021.100033

Piaget, J. (1971). The theory of stages in cognitive development. In D. R. Green, M. P. Ford, & G. B. Flamer (Eds.), *Measurement and Piaget* (pp. 1–11). McGraw-Hill.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and

well-being. *American Psychologist, 55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*(2), 257–285. https://doi.org/10.1207/s15516709cog1202_

Touretzky, D., Gardner-McCune, C., Martin, F., & Seehorn, D. (2019). Envisioning AI for K-12: What should every child know about AI?. *Proceedings of the AAAI Conference on Artificial Intelligence, 33*(01), 9795-9799.

ZawackiRichter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education, 16*, Article 39. https://doi.org/10.1186/s41239-019-0171-0

APPENDICES

APPENDIX: SURVEY QUESTIONS

Language Proficiency

1. On a scale of 1 to 5, how has using AI chatbots improved your language proficiency?

- 1: Not improved at all
- 2: Slightly improved
- 3: Moderately improved
- 4: Significantly improved
- 5: Greatly improved

2. How often do you use AI chatbots to practice your language skills?

- Daily
- Weekly
- Monthly
- Rarely
- Never

3. Which language skills have you found to improve the most with AI chatbot usage?

- Speaking
- Listening
- Reading
- Writing

4. How would you rate the feedback provided by AI chatbots in language learning activities?

- 1: Very poor

- 2: Poor
- 3: Neutral
- 4: Good
- 5: Very good

5. Can you describe a specific instance where an AI chatbot helped you improve your language skills?

6. What features of AI chatbots do you find most helpful for language learning?

7. How do AI chatbots compare to traditional methods in language learning effectiveness?

- 1: Much less effective
- 2: Less effective
- 3: Equally effective
- 4: More effective
- 5: Much more effective

8. How has the use of AI chatbots influenced your confidence in using English?

9. Which AI chatbots have you used the most for language learning? (Please list the names of the chatbots)

10. Which AI chatbot do you find most effective for language learning and why?

Critical Thinking Abilities

11. On a scale of 1 to 5, how have AI chatbots helped improve your critical thinking skills?

- 1: Not at all
- 2: Slightly
- 3: Moderately
- 4: Significantly
- 5: Greatly

12. How often do you engage in activities with AI chatbots that require critical thinking?

- Daily
- Weekly
- Monthly
- Rarely
- Never

13. How effective are AI chatbots in helping you analyze and solve complex problems?h

- 1: Not effective at all
- 2: Slightly effective
- 3: Moderately effective
- 4: Very effective
- 5: Extremely effective

14. Can you provide an example of how an AI chatbot has enhanced your critical thinking abilities?

15. What specific critical thinking skills have you improved through the use of AI chatbots?

- Analysis
- Evaluation
- Inference
- Explanation

16. How do AI chatbots challenge you to think more deeply about course material?

17. In your opinion, how do AI chatbots compare to traditional methods in fostering critical thinking?

- 1: Much less effective
- 2: Less effective

- 3: Equally effective
- 4: More effective
- 5: Much more effective

18. Which AI chatbots have you used the most for developing critical thinking skills? (Please list the names of the chatbots)

19. Which AI chatbot do you find most effective for developing critical thinking skills and why?

Student Engagement

20. On a scale of 1 to 5, how has the use of AI chatbots affected your engagement in learning activities?

- 1: Not at all
- 2: Slightly
- 3: Moderately
- 4: Significantly
- 5: Greatly

21. How often do you use AI chatbots for engaging with course content?

- Daily
- Weekly
- Monthly
- Rarely
- Never

22. How motivating do you find AI chatbots in maintaining your interest in the subject matter?

- 1: Not motivating at all
- 2: Slightly motivating
- 3: Moderately motivating
- 4: Very motivating
- 5: Extremely motivating

23. Can you share a specific example of how an AI chatbot has increased your engagement in a course?

24. What aspects of AI chatbots do you find most engaging?

25. How do AI chatbots support your autonomy in learning?

- 1: Not at all
- 2: Slightly
- 3: Moderately
- 4: Significantly
- 5: Completely

26. How has the use of AI chatbots impacted your interaction with peers and instructors?

27. In what ways do AI chatbots make learning more interactive for you?

28. How do you feel about the balance between AI-driven activities and traditional methods in terms of keeping you engaged?

29. What suggestions do you have for improving the use of AI chatbots to enhance student engagement in your courses?

30. Which AI chatbots have you used the most for increasing engagement? (Please list the names of the chatbots)

31. Which AI chatbot do you find most effective for increasing engagement and why?

Overall

32. Overall, how would you rate the impact of AI chatbots on your learning experience at the Diplomatic Academy of Vietnam?

- 1: Very negative
- 2: Negative
- 3: Neutral
- 4: Positive
- 5: Very positive

Thank you for your participation! Your feedback is crucial in helping us enhance the educational experience at the Diplomatic Academy of Vietnam.