

Teachers' utilization of ChatGPT in teaching elementary English: Basis for the development of guidelines and policies on responsible use of AI in teaching

Luvic Banganan Camalao *

Teaching/Purok 2, Mabiga, Puguin, Conner, Apayao, Philippines
Purok 3, Mabiga, Puguin, Conner, Apayao, Philippines

ABSTRACT

This study investigated the extent of utilization of ChatGPT in teaching elementary English among public-school teachers Southern Conner District, Apayao Province. Using descriptive method research design. The quantitative findings revealed that the overall utilization of ChatGPT was rarely practiced, indicating that its integration into classroom instruction remains minimal. Among the four areas assessed, instructional strategies registered the highest level of use, while assessment and feedback were the least utilized. Significant differences were found when respondents were grouped according to age and familiarity with Generative AI, suggesting that younger and more technologically familiar teachers used ChatGPT more actively. Overall, the study concludes that while ChatGPT presents promising opportunities for language instruction, its adoption among elementary English teachers is still at an early stage, requiring greater digital competence, reliable systems, and structured guidance to ensure its meaningful and responsible use in education. The findings highlight that while ChatGPT has potential as a supportive tool for language instruction, its adoption is still in the early stages due to barriers in training, infrastructure, and policy implementation. The study recommends the establishment of institutional guidelines, continuous capacity-building, and further research on AI literacy and its long-term pedagogical impact in basic education.

Keywords: ChatGPT, Artificial Intelligence, Elementary English, Teaching Strategies, AI Integration, Generative AI, Educational Technology

Date Submitted: April 1, 2026
Date Accepted: April 8, 2026
Date Published: April 14, 2026

INTRODUCTION

Technology has significantly transformed the field of education, particularly in language instruction, where digital tools are increasingly essential for supporting learning. Among these innovations, ChatGPT, a generative artificial intelligence (AI) tool, has introduced new opportunities for automated lesson planning, instant feedback, personalized learning support, and enhanced student engagement. In the context of English language education, ChatGPT provides interactive assistance for reading comprehension, grammar correction, and vocabulary development. However, the success of such integration largely

*Corresponding author/ Email: luvic.camalao@deped.gov.ph
DOI: <http://doi.org/10.69651/PIJHSS0502949>

Recommended citation:

Camalao, L. B. (2026). Teachers' utilization of ChatGPT in teaching elementary English: Basis for the development of guidelines and policies on responsible use of AI in teaching. *Pantao (The International Journal of the Humanities and Social Sciences)* 5 (2), 919-937. <http://doi.org/10.69651/PIJHSS0502949>

depends on how teachers understand, utilize, and adapt ChatGPT in their pedagogical strategies, especially in the elementary level.

The use of ChatGPT in elementary English education presents both opportunities and challenges. While it can support teachers in content development, assessment, and language enrichment, issues related to ethical use, reliability, and accuracy remain significant (Labajova, 2023). Moreover, not all teachers possess the necessary pedagogical skills or technological competence to meaningfully incorporate AI-driven tools into their classrooms. Examining teachers' practices, perceptions, and challenges in using ChatGPT is therefore vital to determine how this tool can enhance student learning while upholding academic integrity. Research on technology integration underscores the importance of teachers' Technological Pedagogical Content Knowledge (TPACK) in ensuring effective adoption (Tseng et al., 2022). Teachers must be able to critically evaluate AI-generated content, align ChatGPT outputs with curriculum standards, and address ethical concerns such as data privacy and authenticity (Le-Nguyen & Tran, 2024; Nedungadi et al., 2024). Institutional support and clear policies also play a crucial role in shaping teachers' readiness and confidence in using ChatGPT as an educational aid.

In the Philippine context, the Department of Education (DepEd) has actively promoted the integration of digital technologies in classrooms through initiatives such as the Digital Rise Program and Learning Management Systems. However, the use of artificial intelligence tools like ChatGPT remains at a nascent stage, with limited guidelines or policies to ensure responsible and effective application. Many educators, particularly in rural areas, continue to face challenges in terms of internet connectivity, access to digital devices, and professional development opportunities in AI literacy. This digital divide restricts equitable adoption of ChatGPT, disadvantaging schools with fewer resources compared to those in more technologically advanced settings (Pattung & Caban, 2023; Kitonga et al., 2024).

Despite its potential, ChatGPT raises critical concerns for education. Questions about the accuracy of its responses, the risks of misinformation, and the ethical implications of using AI-generated materials highlight the need for careful guidance. Over-reliance on ChatGPT may also affect students' critical thinking and problem-solving abilities, making it important to clarify the boundaries of its role in learning. Additionally, teachers' lack of training and preparedness further complicates efforts to integrate ChatGPT into classroom practice (Cabansag, 2022). These challenges are especially evident in underserved areas, where infrastructure and institutional support remain inadequate.

Given these realities, there is a pressing need to explore how teachers currently utilize ChatGPT in teaching elementary English, what benefits and challenges they encounter, and how their experiences can inform the development of guidelines and policies for its responsible use. Understanding these dimensions will enable DepEd and other stakeholders to design targeted interventions, such as teacher training programs, AI literacy workshops, and infrastructure support, ensuring that ChatGPT-enhanced learning is accessible, ethical, and beneficial to all learners regardless of their context.

Statement of the problem

This study was conducted to assess teachers' utilization of ChatGPT in teaching Elementary English as a basis for the development of guidelines and policies on the responsible use of ChatGPT in teaching. Specifically, it sought to examine the characteristics of the respondents, determine the extent of their utilization of ChatGPT across key instructional domains, analyze differences in utilization based on selected profile variables, and formulate appropriate guidelines and policies to ensure its responsible use in educational contexts.

1. What is the profile of the respondents in terms of sex, age, grade level taught, type of school taught, number of years in teaching, highest educational attainment, number of trainings attended related to generative AI, and level of familiarization with the use of generative AI?
2. What is the extent of utilization of ChatGPT by the respondents in teaching Elementary English in terms of instructional strategies, assessment and feedback, student engagement and motivation, and differentiated instruction and personalized learning?
3. Is there a significant difference in the extent of ChatGPT utilization among respondents when grouped according to their profile?
4. What guidelines and policies can be developed to promote the responsible use of ChatGPT in teaching?

METHODOLOGY

This study employed a quantitative research design to determine the extent to which elementary English teachers utilized ChatGPT in their teaching. The design focused on measuring patterns of use across instructional strategies, assessment and feedback, student engagement, and personalized learning activities. The participants of the study were the 48 teachers teaching English in Elementary Schools in Southern Conner District, Apayao Province. To obtain the actual number of respondents, total enumeration was utilized, thereby ensuring that all identified teachers within the specified locale were included in the study.

Data were collected using a structured survey questionnaire designed to capture teachers' utilization of artificial intelligence in instruction. The instrument consisted of two main parts. The first part addressed the demographic profile of the respondents in terms of sex, age, grade level taught, type of school, years of teaching experience, highest educational attainment, AI-related training, and familiarity with generative AI. The second part measured the extent of ChatGPT utilization in teaching and was developed based on existing literature and validated instruments on artificial intelligence integration in education, particularly the works of Rose Luckin and Wayne Holmes. A 5-point Likert scale was employed, ranging from 5 (Always) to 1 (Never), to assess the frequency and extent of teachers' use of ChatGPT across key instructional domains, including instructional strategies, assessment and feedback, student engagement, and personalized learning. To establish content validity, the adapted instrument underwent expert validation by a panel composed of specialists in English language education, educational technology, and artificial intelligence. These experts evaluated the questionnaire in terms of clarity, relevance, and alignment with the research objectives. Their recommendations led to revisions in item wording, removal of ambiguous statements, and refinement of items to ensure appropriateness for elementary English instruction.

Following expert validation, a pilot test was conducted among a group of elementary English teachers who were not included in the final respondents of the study. The purpose of the pilot testing was to assess the clarity of instructions, identify unclear or confusing items, and determine the reliability of the instrument. The data gathered from the pilot test were subjected to reliability analysis using Cronbach's alpha to measure internal consistency. The results yielded acceptable reliability coefficients ($\alpha \geq 0.90$), indicating that the instrument was reliable for measuring the extent of ChatGPT utilization in teaching. Based on the findings of the pilot test and reliability analysis, minor revisions were implemented to further enhance the instrument prior to its final administration.

The data gathering procedure adhered to established ethical research guidelines and was conducted in several phases to ensure the accuracy and dependability of the collected data while maintaining respondent confidentiality. Prior to data collection, the researchers secured

clearance from the Department of Education and other pertinent school authorities, as well as from the appropriate study ethics committee or institutional review board when applicable. Permission was also obtained from participating teachers and school officials, who were requested to provide informed consent outlining the objectives of the study, the confidentiality of data, and the voluntary nature of participation. The survey questionnaire was administered in person to the participants, with printed copies distributed to teachers in schools with limited or no internet connectivity. To improve response rates, reminders were provided during the data collection period, and participants were given one to two weeks to complete the questionnaire at their convenience.

Ethical standards were strictly observed throughout the conduct of the study. Informed consent was obtained from all participants, ensuring that they fully understood the purpose, risks, and benefits associated with the research. Respondents were assured that their participation was voluntary and that they had the right to withdraw at any stage of the study without penalty. Confidentiality and anonymity were maintained, with all collected data securely stored and access restricted solely to authorized researchers. Approval from relevant research ethics bodies was likewise secured to uphold the ethical integrity of the research process.

The analysis of data involved both descriptive and inferential statistical techniques. Frequency and percentage were utilized to describe the demographic profile of the respondents. The weighted mean was employed to assess the extent of respondents' utilization of ChatGPT in teaching. The interpretation of the mean scores was based on the following scale: 4.50 – 5.00 was described as Always and interpreted as ChatGPT is used in all or almost all teaching activities; 3.50 – 4.49 was described as Often and interpreted as ChatGPT is used frequently in teaching activities; 2.50 – 3.49 was described as Sometimes and interpreted as ChatGPT is used occasionally in teaching activities; 1.50 – 2.49 was described as Rarely and interpreted as ChatGPT is seldom used in teaching activities; and 1.00 – 1.49 was described as Never and interpreted as ChatGPT is not used in teaching activities. To determine significant differences in the utilization of generative AI in teaching when respondents were grouped according to profile variables, the Independent Sample T-Test and One Way Analysis of Variance were applied.

RESULTS AND DISCUSSION

The Results and Discussion section presents the findings drawn from the 48 teachers teaching English in Elementary Schools in Southern Conner District, Apayao Province, who served as the respondents of the study through total enumeration. Guided by a quantitative research design, the study examined the extent of teachers' utilization of ChatGPT in teaching Elementary English using a structured survey questionnaire composed of profile variables and indicators covering instructional strategies, assessment and feedback, student engagement and motivation, and differentiated instruction and personalized learning. The data were analyzed through frequency, percentage, weighted mean, Independent Sample T-Test, One Way Analysis of Variance, and post hoc test analysis. The discussion that follows is grounded in the data gathered and interprets the results directly in relation to the objectives of the study, particularly in describing the respondents' profile, determining the extent of ChatGPT utilization across the identified instructional domains, and testing differences in utilization when grouped according to selected profile variables.

The profile of the respondents shows that, in terms of sex, 6 or 12.50% were male and 42 or 87.50% were female. In terms of age, 8 or 16.67% belonged to the 25-30 age group, 14 or 29.17% were 31-35 years old, 12 or 25.00% were 36-40 years old, 6 or 12.50% were 41-45 years old, and 8 or 16.67% were 46 and above. With respect to grade level taught, 5 or 10.42%

were teaching Kindergarten, 5 or 10.42% were teaching Grade 1, 8 or 16.67% were teaching Grade 2, 7 or 14.58% were teaching Grade 3, 3 or 6.25% were teaching Grade 4, 11 or 22.92% were teaching Grade 5, and 9 or 18.75% were teaching Grade 6. In terms of number of years in teaching, 5 or 10.42% had 1-3 years of teaching experience, 6 or 12.50% had 4-6 years, 16 or 33.33% had 7-10 years, 12 or 25.00% had 11-15 years, and 9 or 18.75% had more than 15 years of teaching experience. As to highest educational attainment, 4 or 8.33% held a Bachelor's Degree, 21 or 43.75% had Master's Degree units earned, 19 or 39.58% had completed a Master's Degree, and 4 or 8.33% had Doctorate Degree units earned. In relation to trainings attended related to Gen. AI, 38 or 79.17% had attended none, while 10 or 20.83% had attended trainings. In terms of familiarization with the use of Gen. AI, 11 or 22.92% were not familiar at all, 18 or 37.50% were slightly familiar, 12 or 25.00% were moderately familiar, and 7 or 14.58% were very familiar. These data indicate that the respondents were predominantly female, were mostly within the 31-35 and 36-40 age brackets, were more concentrated in Grade 5 and Grade 6 assignments, had commonly accumulated 7-10 years of teaching experience, and generally possessed graduate level academic advancement through either earned or completed Master's Degree work. At the same time, the profile also reveals that 79.17% had not attended trainings related to Gen. AI, while only 20.83% had done so, and that most respondents were only slightly familiar or moderately familiar with Generative AI. This demographic and professional context is important because it situates the relatively limited exposure of the respondents to formal AI capacity building, which may help explain the subsequent patterns in ChatGPT utilization.

In the area of instructional strategies, the extent of utilization of ChatGPT in teaching Elementary English yielded an average mean of 2.19, described as "Rarely." The individual item means further showed that teachers used ChatGPT to create lesson plans and instructional materials for teaching English with a mean of 2.02, described as "Rarely"; incorporated ChatGPT-generated content such as stories and dialogues into lessons with a mean of 2.19, described as "Rarely"; utilized ChatGPT to provide real-time explanations of grammar and vocabulary with a mean of 2.44, described as "Rarely"; adapted ChatGPT-generated reading passages to suit different student proficiency levels with a mean of 2.15, described as "Rarely"; used ChatGPT to enhance storytelling and interactive teaching strategies with a mean of 2.23, described as "Rarely"; leveraged ChatGPT-generated exercises and drills to reinforce English language concepts with a mean of 2.29, described as "Rarely"; integrated ChatGPT-generated visual or textual aids to support English instruction with a mean of 1.98, described as "Rarely"; used ChatGPT to generate comprehension questions and activities for texts with a mean of 2.19, described as "Rarely"; modified ChatGPT-generated lesson suggestions to align with the curriculum with a mean of 2.40, described as "Rarely"; and relied on ChatGPT to translate or simplify English texts for students when needed with a mean of 2.00, described as "Rarely." The data show that the highest mean score in this domain was 2.44 for providing real-time explanations of grammar and vocabulary, followed by 2.40 for modifying ChatGPT-generated lesson suggestions to align with the curriculum and 2.29 for leveraging ChatGPT-generated exercises and drills to reinforce English language concepts. Conversely, the lowest mean scores were 1.98 for integrating ChatGPT-generated visual or textual aids, 2.00 for translating or simplifying English texts, and 2.02 for creating lesson plans and instructional materials. Although all indicators were interpreted uniformly as "Rarely," the distribution of means suggests that teachers were somewhat more inclined to use ChatGPT for direct language clarification, curriculum adjustment, and practice activities than for material production, translation support, or visual and textual aid generation. The overall mean of 2.19 confirms that ChatGPT was seldom integrated into instructional strategies and had not yet become a regular component of classroom pedagogy in teaching Elementary English.

This low level of utilization in instructional strategies indicates that the integration of ChatGPT in classroom teaching remained minimal and selective. The pattern suggests that teachers viewed ChatGPT more as a supplementary support tool for language explanation and reinforcement than as a central instructional resource. Such findings are consistent with the observation of Qadir (2023), who emphasized that teachers in basic education are often hesitant to adopt AI tools because of concerns about accuracy, curriculum alignment, and the complexity of emerging technologies. The result also aligns with Chang and Kidman (2023), who found that in the absence of clear school policies and sustained professional development, teachers limit their use of ChatGPT to low-risk or superficial tasks. The comparatively higher means for grammar and vocabulary explanations, curriculum-aligned lesson modification, and exercise generation indicate that teachers were beginning to explore functions of ChatGPT that are closely tied to immediate instructional needs. This supports the view of Song et al. (2025), who noted that ChatGPT can assist lesson preparation by providing readily accessible instructional materials and exemplars that may be adapted to learner needs. Similarly, Borrás et al. (2024) observed that teachers increasingly use generative AI for idea generation, vocabulary support, and text simplification, particularly when responding to diverse student needs. In the same way, Moorhouse and Kohnke (2024) reported that English teachers see ChatGPT as a useful scaffolding tool for producing varied learning materials more efficiently, while Ooi et al. (2023) argued that ChatGPT functions most effectively as a pedagogical assistant rather than a pedagogical replacement. Verma and Prasad (2023) also found that teachers show greater confidence in integrating ChatGPT for content enhancement than for direct instructional delivery. Taken together, the instructional strategies results suggest that while respondents recognized some practical value in ChatGPT, they had not yet incorporated it into a broader, more systematic pedagogical framework.

In the domain of assessment and feedback, the respondents registered an overall mean of 1.78, with a descriptive equivalent of “Never.” The item means showed that teachers used ChatGPT to create quizzes and tests for English assessment with a mean of 1.92, described as “Rarely”; leveraged ChatGPT to generate immediate feedback for student responses with a mean of 1.77, described as “Never”; utilized ChatGPT-powered grammar and spelling support to assist students in writing tasks with a mean of 1.90, described as “Rarely”; employed ChatGPT to generate rubrics and scoring criteria for English assignments with a mean of 1.96, described as “Rarely”; used ChatGPT-based tools to track student progress in language learning with a mean of 1.79, described as “Never”; relied on ChatGPT to generate alternative test questions to reduce academic dishonesty with a mean of 1.67, described as “Never”; used ChatGPT to provide personalized feedback on students’ spoken and written English with a mean of 1.67, described as “Never”; incorporated ChatGPT in peer assessment activities to enhance feedback accuracy with a mean of 1.65, described as “Never”; utilized ChatGPT-powered text analysis to identify common language errors in student work with a mean of 1.75, described as “Never”; and used ChatGPT-generated feedback to adjust instructional strategies with a mean of 1.71, described as “Never.” The highest mean in this domain was 1.96 for generating rubrics and scoring criteria, followed by 1.92 for creating quizzes and tests, and 1.90 for grammar and spelling support in writing tasks. The lowest mean was 1.65 for incorporating ChatGPT in peer assessment activities, while both personalized feedback on spoken and written English and generating alternative test questions to reduce academic dishonesty each posted a mean of 1.67. Since the overall mean of 1.78 is interpreted as “Never,” the findings reveal that teachers did not commonly use ChatGPT for assessment and feedback purposes. Even the highest scoring indicators remained within low utilization levels, showing that assessment-related functions of ChatGPT were largely absent from regular instructional practice.

The minimal integration of ChatGPT in assessment and feedback indicates that teachers remained cautious in using AI for evaluative functions, particularly where judgment, feedback accuracy, and student performance monitoring are involved. The pattern suggests that respondents were more comfortable using ChatGPT for basic assessment preparation tasks, such as rubric development and quiz creation, than for dynamic and personalized feedback processes. This is consistent with Han et al. (2024), who found that while AI-powered feedback can improve young learners' writing accuracy and self-correction skills, many teachers hesitate to rely on automated feedback tools due to doubts about the accuracy and appropriateness of AI-generated corrections. In a similar vein, Anis (2024) noted that teachers in elementary English settings continue to depend on traditional assessment methods despite recognizing the potential of ChatGPT for grammar checking, text analysis, and writing development. Hwang and Chen (2023) likewise emphasized that AI-driven evaluation systems require teacher confidence and technical competence that are not yet fully developed in many school environments. Their findings support the low means observed in student progress tracking, alternative test item generation, personalized feedback, peer assessment, and instructional adjustment through AI-generated feedback. Tai and Chen (2024) further noted that AI-generated instructional and assessment materials may improve comprehension and performance, yet teachers often struggle to integrate these functions meaningfully due to limited AI literacy. El-Seoud et al. (2023) also reported that concerns about ethical use, content accuracy, and bias in AI-generated outputs hinder fuller adoption of AI-based assessment tools. The present findings resonate strongly with those concerns, as respondents rarely or never used ChatGPT for more advanced, interactive, or student-centered assessment tasks. Boras et al. (2024) and Jauhiainen and Guerra (2023) likewise observed that AI is often limited to basic text assistance rather than higher-order formative or interactive assessment. Overall, the overall mean of 1.78 demonstrates that ChatGPT had minimal influence on the assessment and feedback practices of the respondents, making this the weakest among all identified domains of utilization.

With respect to student engagement and motivation, the overall mean was 1.86, described as "Rarely," indicating that teachers seldom used ChatGPT to make English classes more engaging and motivating. The item means revealed that respondents used ChatGPT-generated stories and interactive dialogues to make English lessons more engaging with a mean of 2.08, described as "Rarely"; integrated ChatGPT-based chatbots to encourage students to practice conversational English with a mean of 1.71, described as "Never"; utilized ChatGPT-generated gamified exercises to sustain student interest in learning English with a mean of 1.94, described as "Rarely"; employed ChatGPT-assisted speech or pronunciation tools to help students improve spoken English with a mean of 1.85, described as "Rarely"; used ChatGPT-generated role-playing scenarios to enhance student participation with a mean of 1.81, described as "Rarely"; leveraged ChatGPT-generated visual or audio content to cater to different learning styles with a mean of 1.83, described as "Rarely"; encouraged students to interact with ChatGPT-driven platforms to reinforce English skills with a mean of 1.75, described as "Never"; integrated ChatGPT-generated creative prompts such as humor, storytelling, and writing to spark student creativity with a mean of 1.96, described as "Rarely"; utilized ChatGPT-powered digital assistants to answer students' language-related questions with a mean of 1.77, described as "Never"; and incorporated ChatGPT tools in group activities to promote collaboration and active learning with a mean of 1.85, described as "Rarely." The highest mean of 2.08 was recorded for using ChatGPT-generated stories and interactive dialogues, followed by 1.96 for creative prompts and 1.94 for gamified exercises. The lowest means were 1.71 for chatbot integration to encourage conversational English, 1.75 for encouraging student interaction with ChatGPT-driven platforms, and 1.77 for utilizing

ChatGPT-powered digital assistants to answer language-related questions. These results show that while some teachers occasionally used ChatGPT to generate creative and engaging materials, the more interactive and student-driven uses of ChatGPT were almost never practiced. The overall mean of 1.86 therefore indicates that ChatGPT had only limited integration in strategies intended to promote engagement, motivation, and participatory learning in Elementary English.

The results in this domain suggest that teachers confined their use of ChatGPT mainly to passive content enhancement rather than direct student interaction with AI-supported tools. This means that the technology was seldom employed as a catalyst for participatory, conversational, or autonomous engagement, despite its potential to support such learning experiences. The findings are consistent with Boras et al. (2024), who found that although students frequently use AI tools for translation, explanation, and idea generation, teachers rarely design AI-driven interactive tasks that produce authentic engagement. Jauhiainen and Guerra (2023) similarly emphasized that AI-based conversational tools can create low-pressure environments for practicing language skills, yet teachers often underutilize these features because of concerns related to accuracy, classroom management, and unfamiliarity with AI-supported pedagogy. The relatively higher means for stories, creative prompts, and gamified exercises indicate that respondents recognized some value in ChatGPT for generating materials that could capture student interest. This aligns with Tai and Chen (2024), who reported that generative AI can enhance motivation through personalized passages, interactive storytelling, and creative prompts, but that many teachers still lack sufficient training to implement these effectively. Hiniz (2024) also argued that ChatGPT can improve inclusivity and foster learner participation in EFL classrooms when teachers intentionally design opportunities for students to engage with AI-generated content. However, the low means for chatbot-based conversational practice, student interaction with ChatGPT-driven platforms, and digital assistant use suggest that such pedagogically richer possibilities remained largely unrealized among the respondents. Mao et al. (2024) further demonstrated that generative AI can support gamified and immersive learning environments, but these applications require technological support and pedagogical confidence that many basic education teachers do not yet possess. Anis (2024) likewise noted that teachers often return to traditional methods because of concerns about students' overreliance on AI tools. The present findings therefore suggest that while ChatGPT had some limited use in motivating learners through generated content, its role in fostering interactive, participatory, and student-centered engagement remained minimal.

In the area of differentiated instruction and personalized learning, the respondents obtained an overall mean of 1.94, described as "Rarely," which indicates that teachers seldom used ChatGPT to support learner differences and individualized instruction. The item means showed that teachers used ChatGPT to customize lesson plans based on student proficiency levels with a mean of 2.06, described as "Rarely"; employed ChatGPT-powered tools to provide individualized learning pathways with a mean of 1.90, described as "Rarely"; leveraged ChatGPT-generated adaptive quizzes to assess students at their own pace with a mean of 1.98, described as "Rarely"; utilized ChatGPT tools to generate differentiated exercises based on student needs with a mean of 2.27, described as "Rarely"; used ChatGPT-generated reading materials tailored to different skill levels with a mean of 2.15, described as "Rarely"; relied on ChatGPT tools to provide personalized writing prompts for students with a mean of 1.73, described as "Never"; integrated ChatGPT-driven language tutors to offer individualized support to students with a mean of 1.79, described as "Never"; utilized ChatGPT-generated feedback to tailor instruction for struggling students with a mean of 1.81, described as "Rarely"; used ChatGPT to recommend supplementary resources based on students' performance with a mean of 1.88, described as "Rarely"; and implemented ChatGPT-powered text-to-speech or speech-to-text tools for diverse learners with a mean of 1.79,

described as “Never.” The highest mean in this domain was 2.27 for generating differentiated exercises based on student needs, followed by 2.15 for generating reading materials tailored to different skill levels and 2.06 for customizing lesson plans based on student proficiency levels. The lowest mean was 1.73 for providing personalized writing prompts, while both integrating ChatGPT-driven language tutors and implementing text-to-speech or speech-to-text tools each obtained a mean of 1.79. These data show that teachers occasionally used ChatGPT to adjust exercises, reading materials, and lesson plans to student levels, but did not commonly adopt more individualized, assistive, or adaptive applications. The overall mean of 1.94 confirms that the role of ChatGPT in differentiated instruction and personalized learning remained minimal and infrequent.

This limited use of ChatGPT for differentiation and personalization implies that the respondents had not yet maximized the potential of AI for tailoring instruction to individual learner needs. The findings suggest that ChatGPT was used only in modest ways to modify content difficulty and learning tasks, while more sophisticated applications such as individualized learning pathways, AI tutoring, personalized feedback, and accessibility features were rarely or never implemented. This pattern is consistent with Hwang and Chen (2023), who emphasized that AI-enhanced personalized learning depends on technological competence, pedagogical readiness, and institutional support, all of which are still developing among many teachers in basic education. Le-Nguyen and Tran (2024) likewise observed that although ChatGPT can generate leveled texts, scaffold writing tasks, and adjust linguistic complexity, many teachers lack the confidence and training to embed these adaptive features into regular classroom practice. Chounta et al. (2021) similarly pointed out that while generative AI can scaffold learning materials and modify tasks based on learner needs, such features are often underutilized because teachers are unfamiliar with AI-driven differentiation frameworks. Moorhouse and Kohnke (2024) further found that English teachers generally use ChatGPT for content enhancement rather than targeted personalization, which corresponds closely with the present results. Chan and Tsi (2024) also argued that without strong technological pedagogical content knowledge, tools such as ChatGPT remain supplementary rather than integral to personalized instruction. Kasneci et al. (2023) reinforced this point by asserting that personalized learning through AI requires reliable digital infrastructure and ongoing professional development, both of which remain inconsistent across school settings. Tai and Chen (2024) likewise observed that AI-supported personalization can improve reading comprehension and writing performance, but its implementation is constrained by limited AI literacy among teachers. Anis (2024) additionally cautioned that in the absence of clear guidelines, teachers may be reluctant to depend on AI for individualized support due to concerns about accuracy, appropriateness, and student overdependence. Zhai (2023) emphasized that meaningful integration of AI into personalized learning requires long-term capacity building rather than isolated training sessions. The overall mean of 1.94 therefore reveals that ChatGPT’s potential to support differentiated and personalized instruction remained largely untapped among the respondents.

When the four domains were summarized, the total average weighted mean on the extent of utilization of ChatGPT in teaching Elementary English was 1.94, described as “Rarely.” Among the areas, Instructional Strategies obtained the highest average mean of 2.19, described as “Rarely,” followed by Differentiated Instruction and Personalized Learning with 1.94, also described as “Rarely,” and Student Engagement and Motivation with 1.86, described as “Rarely.” The lowest average mean was recorded in Assessment and Feedback with 1.78, described as “Never.” These results indicate that although all domains reflected generally low levels of ChatGPT utilization, teachers were relatively more inclined to use the tool for lesson-related and content-adjustment purposes than for evaluation, monitoring, feedback, and

interactive student engagement. The total average weighted mean of 1.94 confirms that ChatGPT was seldom employed in the teaching practices of the respondents and had not yet become a regular component of Elementary English instruction. The pattern across domains suggests a cautious and emerging form of AI integration, in which teachers utilized ChatGPT only in areas perceived as manageable, low risk, and immediately supportive of lesson preparation. The lowest score in assessment and feedback further implies that teachers exercised the greatest caution in pedagogical areas requiring professional judgment, ethical sensitivity, and direct accountability for student evaluation. Overall, the data demonstrate that ChatGPT remained on the periphery of classroom practice rather than functioning as a fully integrated instructional technology.

The generally low utilization across all four domains underscores that AI integration in this context is still at an early stage of development. Even in the highest domain, Instructional Strategies, the average mean was only 2.19 and remained within the "Rarely" category, while Assessment and Feedback registered 1.78 and was categorized as "Never." This pattern is consistent with current scholarship showing that teachers tend to begin with practical, content-oriented uses of generative AI before progressing to more complex applications involving personalization, assessment, or autonomous student interaction. The present findings also reflect broader issues related to teacher readiness, policy absence, and limited professional development. In this context, the low overall utilization can be interpreted not simply as resistance, but also as evidence of a transitional phase in which respondents were aware of ChatGPT and occasionally explored it, yet lacked the training, support systems, and institutional guidance needed for more confident and sustained adoption. This has practical implications for school leadership and policy makers because the results point to the need for structured guidelines, clearer governance on responsible AI use, and targeted professional development that addresses both technical skills and pedagogical application.

The test of difference on the extent of utilization of ChatGPT in teaching Elementary English when respondents were grouped according to their profile variables revealed mixed results. Using the decision rule that if $p\text{-value} < 0.05$, then reject H_0 , the findings showed that sex had a $p\text{-value}$ of 0.530 and was interpreted as Not Significant, age had a $p\text{-value}$ of 0.020 and was interpreted as Significant, grade level taught had a $p\text{-value}$ of 0.681 and was interpreted as Not Significant, number of years in teaching had a $p\text{-value}$ of 0.111 and was interpreted as Not Significant, highest educational attainment had a $p\text{-value}$ of 0.494 and was interpreted as Not Significant, trainings attended had a $p\text{-value}$ of 0.134 and was interpreted as Not Significant, and familiarization with the use of Gen. AI had a $p\text{-value}$ of 0.000 and was interpreted as Significant. Based on the stated criterion $p\text{-value} < 0.05$, the null hypothesis was rejected for age and for familiarization with the use of Gen. AI, and the null hypothesis was not rejected for sex, grade level taught, number of years in teaching, highest educational attainment, and trainings attended. These results indicate that teachers differed significantly in their extent of ChatGPT utilization when grouped according to age and familiarization with Generative AI, but did not differ significantly when grouped according to sex, grade level taught, years in teaching, educational attainment, or training attendance. The significance of age and familiarization implies that the extent of ChatGPT utilization was meaningfully influenced by generational and technological familiarity factors, while the lack of significant differences in the remaining variables suggests that ChatGPT use was broadly similar across those categories.

The significant difference according to age indicates that age functioned as an influential variable in shaping ChatGPT utilization. This result suggests that respondents' readiness or inclination to use ChatGPT was not evenly distributed across age groups. The finding is consistent with Anis and Khalid (2024), who observed that younger educators are generally more open to adopting innovative technologies because of greater digital literacy and

wider exposure to emerging tools. In the same way, Chan and Tsi (2024) described younger educators as “digital natives” who are more naturally attuned to technology-based tools, whereas older teachers, often referred to as “digital immigrants,” may require more time and support to adapt to digital pedagogies. At the same time, the absence of significant differences by sex, grade level taught, number of years in teaching, highest educational attainment, and trainings attended suggests that these variables did not substantially differentiate respondents’ use of ChatGPT. The p-values of 0.530, 0.681, 0.111, 0.494, and 0.134, all exceeding 0.05, indicate that any observed differences across these categories were not statistically significant. In methodological terms, the null hypothesis was not rejected for these variables. This means that male and female teachers, teachers across grade levels, teachers with varying years of experience, teachers with different educational qualifications, and teachers who had or had not attended trainings related to Generative AI did not significantly differ in their extent of ChatGPT utilization. The lack of significance for trainings attended is particularly notable, because although one might expect formal exposure to training to influence utilization, the result suggests that the existing trainings may not have been sufficient in depth, frequency, or practical application to produce statistically meaningful differences in actual classroom use.

The post hoc test analysis on the extent of utilization of ChatGPT when grouped according to age clarified where the significant age-based differences occurred. The descriptive statistics showed that teachers aged 25-30 years old obtained the highest mean utilization score of $M = 2.72$, $SD = 0.966$, followed by those aged 31-35 years old with $M = 2.14$, $SD = 0.970$, and those aged 41-45 years old with $M = 2.02$, $SD = 0.852$. The lowest mean scores were recorded among teachers aged 36-40 years old with $M = 1.47$, $SD = 0.565$ and those aged 46 years and above with $M = 1.46$, $SD = 0.446$. The post hoc comparison matrix showed that the mean difference between 31-35 yrs old and 36-40 yrs old was 0.677 with p-value = 0.214, between 31-35 yrs old and 46 and above was 0.68411 with p-value = 0.202, between 31-35 yrs old and 41-45 yrs old was 0.123 with p-value = 0.998, and between 31-35 yrs old and 25-30 yrs old was -0.581 with p-value = 0.663. The mean difference between 36-40 yrs old and 46 and above was 0.00708 with p-value = 1.000, between 36-40 yrs old and 41-45 yrs old was -0.554 with p-value = 0.623, and between 36-40 yrs old and 25-30 yrs old was -1.258* with p-value = 0.046. The mean difference between 46 and above and 41-45 yrs old was -0.561 with p-value = 0.609, and between 46 and above and 25-30 yrs old was -1.265* with p-value = 0.045. The mean difference between 41-45 yrs old and 25-30 yrs old was -0.704 with p-value = 0.614. The notation indicated that * $p < .05$, ** $p < .01$, and *** $p < .001$. Based on these exact results, statistically significant differences were found only between teachers aged 25-30 years old and those aged 36-40 years old, where $p = 0.046$, and between teachers aged 25-30 years old and those aged 46 years and above, where $p = 0.045$. Since both p-values are less than 0.05, the null hypothesis for these specific pairwise comparisons was rejected. All other pairwise comparisons had p-values greater than 0.05, so the null hypothesis was not rejected for those age-group pairs.

These post hoc findings show that the significant age effect identified in the overall test was driven primarily by the markedly higher ChatGPT utilization of the youngest group, those aged 25-30 years old, compared with the 36-40 years old group and the 46 years and above group. The mean scores of $M = 2.72$ for 25-30 years old respondents, $M = 1.47$ for 36-40 years old respondents, and $M = 1.46$ for respondents aged 46 years and above demonstrate a clear gradient in utilization, with younger teachers showing greater use of ChatGPT in teaching Elementary English. This pattern may be explained by differences in confidence, familiarity, and openness to digital innovation across age groups. Dwivedi et al. (2023), drawing on the Technology Acceptance Model (TAM), explained that younger teachers are more likely to perceive technology as useful and easy to use, which increases their behavioral intention to

adopt it. In contrast, older teachers may have more established routines and may perceive AI tools as more complex or less necessary in their instructional practice. This interpretation is also consistent with Kitonga et al. (2024), who emphasized that teachers' digital competence and attitudes toward technology are shaped by generational exposure as well as experience. The present results therefore indicate that the significant difference in ChatGPT utilization according to age was not general across all age groups but was specifically attributable to the greater utilization among teachers aged 25-30 years old in comparison with selected older groups.

The post hoc test analysis on the extent of utilization of ChatGPT when grouped according to respondents' familiarization with the use of AI also clarified the source of the significant difference found in the overall test. The descriptive statistics showed that teachers who were very familiar with AI obtained the highest mean utilization score of $M = 2.71$, $SD = 0.800$, followed by those who were moderately familiar with $M = 2.44$, $SD = 0.890$. Teachers who were slightly familiar had a mean of $M = 1.77$, $SD = 0.651$, while those who were not familiar at all had the lowest mean score of $M = 1.18$, $SD = 0.597$. The post hoc comparison table showed that the mean difference between Moderately Familiar and Slightly Familiar was 0.669 with p -value = 0.150, between Moderately Familiar and Very Familiar was -0.273 with p -value = 0.899, and between Moderately Familiar and Not Familiar at all was 1.261** with p -value = 0.004. The mean difference between Slightly Familiar and Very Familiar was -0.942 with p -value = 0.082, and between Slightly Familiar and Not Familiar at all was 0.592 with p -value = 0.086. The mean difference between Very Familiar and Not Familiar at all was 1.534** with p -value = 0.006. The notation again indicated that * $p < .05$, ** $p < .01$, and *** $p < .001$. These results show that statistically significant differences occurred between teachers who were Moderately Familiar and those who were Not Familiar at all, where $p = 0.004$, and between teachers who were Very Familiar and those who were Not Familiar at all, where $p = 0.006$. Since both p -values are less than 0.05, the null hypothesis for these pairwise comparisons were rejected. No significant differences were found among the other pairs because their p -values of 0.150, 0.899, 0.082, and 0.086 all exceeded 0.05, so the null hypothesis was not rejected for those comparisons.

These findings clearly indicate that familiarity with AI was a major factor influencing teachers' use of ChatGPT. The progression in mean scores from $M = 1.18$, $SD = 0.597$ among those not familiar at all, to $M = 1.77$, $SD = 0.651$ among those slightly familiar, to $M = 2.44$, $SD = 0.890$ among those moderately familiar, and finally to $M = 2.71$, $SD = 0.800$ among those very familiar, demonstrates an increasing trend in ChatGPT utilization as familiarity with AI improves. The significant differences between the Moderately Familiar and Not Familiar at all groups and between the Very Familiar and Not Familiar at all groups confirm that teachers with higher levels of AI familiarization used ChatGPT significantly more than those with no familiarity. This supports the argument of Alasadi and Baiz (2023), who found that prior exposure to AI tools enhances teachers' capacity to integrate such technologies effectively into lesson planning and instructional delivery. Chang and Kidman (2023) similarly noted that technological familiarity directly influences self-efficacy in digital pedagogy, thereby shaping both the extent and quality of technology integration. The pattern also aligns closely with the Technological Pedagogical Content Knowledge framework of Mishra and Koehler (2006), which emphasizes that technological knowledge is an essential foundation for effective technology integration. Teachers with greater familiarization are more capable of balancing technology, pedagogy, and content, and are therefore better positioned to use ChatGPT meaningfully in their teaching. In contrast, respondents who were not familiar at all, with $M = 1.18$ and $SD = 0.597$, demonstrated the lowest utilization and differed significantly from the moderately familiar and very familiar groups. This suggests that lack of technological familiarity may severely limit a teacher's readiness to integrate Generative AI into instructional

design, assessment, or differentiated learning. The practical implication is that familiarity is not merely a background characteristic but a decisive condition for adoption.

Overall, the results and discussion demonstrate that the 48 respondents exhibited generally low utilization of ChatGPT in teaching Elementary English, with a total average weighted mean of 1.94, interpreted as “Rarely.” Across the four domains, Instructional Strategies obtained the highest average mean of 2.19, Differentiated Instruction and Personalized Learning obtained 1.94, Student Engagement and Motivation obtained 1.86, and Assessment and Feedback obtained the lowest mean of 1.78, interpreted as “Never.” These findings indicate that ChatGPT was used sparingly and mainly in lesson-related functions rather than in assessment, interactive engagement, or personalized support. The profile data further show that most respondents were female, many were in the 31-35 and 36-40 age groups, most had 7-10 years of teaching experience, many had Master’s Degree units earned or completed, 79.17% had no training related to Gen. AI, and most were only slightly familiar or moderately familiar with Generative AI. Inferential analysis revealed that age, with $p = 0.020$, and familiarization with the use of Gen. AI, with $p = 0.000$, significantly influenced the extent of ChatGPT utilization, leading to the rejection of the null hypothesis for these variables, while sex, grade level taught, number of years in teaching, highest educational attainment, and trainings attended did not significantly influence utilization, leading to the non-rejection of the null hypothesis for those variables. The post hoc analyses further established that significant age-based differences existed between respondents aged 25-30 years old and those aged 36-40 years old, $p = 0.046$, and between respondents aged 25-30 years old and those aged 46 years and above, $p = 0.045$, while significant familiarity-based differences existed between Moderately Familiar and Not Familiar at all, $p = 0.004$, and between Very Familiar and Not Familiar at all, $p = 0.006$. Collectively, these findings make a clear contribution to the understanding of AI integration in basic education by showing that ChatGPT use among elementary English teachers remains in an early and cautious stage, and that younger age and stronger AI familiarization are the most meaningful factors associated with increased utilization. These results directly address the objectives of the study and provide a strong empirical basis for the subsequent development of guidelines and policies on the responsible use of ChatGPT in teaching, as well as for future recommendations on training, institutional support, and responsible pedagogical integration of Generative AI in elementary education.

CONCLUSION

The findings of the study demonstrate that the utilization of ChatGPT in teaching Elementary English among the 48 respondents remains generally minimal, as reflected in the consistently low levels of use across instructional strategies, assessment and feedback, student engagement and motivation, and differentiated instruction and personalized learning. The overall average weighted mean of 1.94, interpreted as “Rarely,” confirms that ChatGPT has not yet become an integral component of classroom practice. Among the four domains, instructional strategies registered the highest average mean of 2.19, interpreted as “Rarely,” indicating that teachers were more inclined to use ChatGPT for lesson-related functions such as grammar explanation, content modification, and drill generation. This was followed by differentiated instruction and personalized learning with an average mean of 1.94, and student engagement and motivation with 1.86, both also interpreted as “Rarely.” The lowest utilization was observed in assessment and feedback, with an average mean of 1.78, interpreted as “Never,” demonstrating that ChatGPT was least applied in evaluative and feedback-oriented processes. These results collectively indicate that while some exploratory use of ChatGPT

exists, particularly in content support and lesson preparation, its application in more complex pedagogical domains remains limited and underdeveloped.

The demographic and professional profile of the respondents provides further context for understanding these findings. The majority of the teachers were female, with 42 or 87.50% compared to 6 or 12.50% male respondents, and most were within the 31-35 age group at 29.17% and 36-40 age group at 25.00%. In terms of grade level, a larger proportion handled Grade 5 at 22.92% and Grade 6 at 18.75%, while teaching experience was most concentrated among those with 7-10 years at 33.33% and 11-15 years at 25.00%. Most respondents had advanced academic qualifications, with 43.75% having Master's Degree units earned and 39.58% having completed a Master's Degree, yet only 8.33% had Doctorate Degree units earned and 8.33% held only a Bachelor's Degree. Despite this relatively high level of academic preparation, 79.17% of respondents had not attended any training related to Generative AI, and only 20.83% had participated in such training. In terms of familiarity, 37.50% were slightly familiar, 25.00% were moderately familiar, 22.92% were not familiar at all, and only 14.58% were very familiar. These figures suggest that although the respondents possess substantial teaching experience and academic advancement, their exposure to AI-related training and their familiarity with Generative AI remain limited, which likely contributed to the low levels of ChatGPT utilization observed in the study.

The inferential analysis further strengthens this interpretation by revealing that significant differences in ChatGPT utilization were found only when respondents were grouped according to age and familiarization with the use of Generative AI. The p-value for age was 0.020, and the p-value for familiarization was 0.000, both of which are less than the 0.05 level of significance, leading to the rejection of the null hypothesis for these variables. In contrast, no significant differences were observed when respondents were grouped according to sex, with $p = 0.530$, grade level taught, with $p = 0.681$, number of years in teaching, with $p = 0.111$, highest educational attainment, with $p = 0.494$, and trainings attended related to Generative AI, with $p = 0.134$, as all these values exceeded the 0.05 level of significance, resulting in the non-rejection of the null hypothesis. The post hoc analysis clarified that teachers aged 25-30 years old had significantly higher utilization of ChatGPT, with $M = 2.72$, $SD = 0.966$, compared to those aged 36-40 years old, with $M = 1.47$, $SD = 0.565$, where $p = 0.046$, and those aged 46 years and above, with $M = 1.46$, $SD = 0.446$, where $p = 0.045$. Similarly, teachers who were very familiar with AI, with $M = 2.71$, $SD = 0.800$, and those who were moderately familiar, with $M = 2.44$, $SD = 0.890$, demonstrated significantly higher utilization than those who were not familiar at all, with $M = 1.18$, $SD = 0.597$, where $p = 0.006$ and $p = 0.004$, respectively. These findings indicate that both generational factors and technological familiarity play a critical role in influencing the adoption and use of ChatGPT in teaching, while other demographic and professional variables appear to have no statistically significant effect.

Taken together, the results confirm that although ChatGPT presents promising opportunities to enhance language instruction through content generation, feedback support, engagement strategies, and personalized learning, its actual classroom integration among elementary English teachers remains at an early and cautious stage. The limited use across all domains reflects a combination of factors, including insufficient training, low levels of familiarity, concerns about accuracy and appropriateness, and the absence of clear institutional guidelines for AI integration. These findings underscore the need for a more systematic and structured approach to the adoption of ChatGPT in education, one that not only introduces teachers to the technical aspects of AI tools but also equips them with the pedagogical competencies necessary to integrate these tools effectively and responsibly.

In light of these conclusions, it becomes imperative for the Department of Education to implement sustained training and capacity-building programs that focus on the pedagogical application of ChatGPT and other AI technologies. Such programs should extend beyond basic

technical orientation and emphasize classroom integration, content evaluation, and ethical considerations in AI use. At the same time, schools should establish clear and comprehensive policies that define the appropriate use of ChatGPT, set boundaries for both teachers and students, ensure data privacy, and uphold academic integrity while encouraging innovation in instructional practices. Furthermore, educational institutions, particularly those in rural areas, must invest in improving internet connectivity, providing access to digital devices, and strengthening technical support systems to ensure that AI-assisted instruction becomes both feasible and sustainable.

Teachers are likewise encouraged to adopt ChatGPT as a complementary instructional tool rather than as a replacement for traditional teaching practices. Its use should be directed toward enhancing creativity, collaboration, and critical thinking among students, while maintaining the teacher's central role as a facilitator of learning. Finally, future research is necessary to further explore the long-term effects of ChatGPT-assisted instruction on students' language proficiency, motivation, and learning outcomes. Subsequent studies may also examine the evolving perceptions and practices of teachers as AI tools become more accessible, conduct comparative analyses across educational levels and subject areas, and investigate the ethical and policy implications of AI integration in basic education. Through these combined efforts, the meaningful, equitable, and responsible use of ChatGPT and other Generative AI technologies in Philippine classrooms can be progressively realized.

REFERENCES

- Abunaseer, H. (2023). The use of generative AI in education: Applications, and impact. *Technology and the Curriculum: Summer 2023*.
- Alasadi, E. A., & Baiz, C. R. (2023). Generative AI in education and research: Opportunities, concerns, and solutions. *Journal of Chemical Education*, 100(8), 2965-2971. <https://doi.org/10.1021/acs.jchemed.3c00323>
- Anis, N. A. M. O. U. S., & Khalid, B. O. U. I. G. H. A. (2024). Teachers' and students' perception about the effects of using ChatGPT, Grammarly, and Quillbot on students' writing skills (Doctoral dissertation, University Center of Abdelhafid Boussouf-Mila).
- Baidoo-Anu, D., & Owusu Ansah, L. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, 7(1), 52-62. <https://doi.org/10.61969/jai.1337500>
- Barrett, A., & Pack, A. (2023). Not quite eye to AI: Student and teacher perspectives on the use of generative artificial intelligence in the writing process. *International Journal of Educational Technology in Higher Education*, 20(1). <https://doi.org/10.1186/s41239-023-00427-0>
- Boras, B., Smolić, E., Gledec, G., & Jaguš, T. (2024). Exploring the educational potential of generative AI: An application for spelling practice. In *INTED2024 proceedings* (pp. 6945-6952). IATED.
- Bozkurt, A. (2023). ChatGPT, üretken yapay zeka ve algoritmik paradigma değişikliği. *Alanyazın*, 4(1), 63-72. <https://doi.org/10.59320/alanyazin.1283282>

- Cabansag, J. N. (2022). Lived experiences of higher education institution students in online learning classes in the new normal. *World Journal of English Language*, 12(7), 265-265.
- Chan, C. K., & Tsi, L. H. (2024). Will generative AI replace teachers in higher education? A study of teacher and student perceptions. *Studies in Educational Evaluation*, 83, 101395. <https://doi.org/10.1016/j.stueduc.2024.101395>
- Chang, C., & Kidman, G. (2023). The rise of generative artificial intelligence (AI) language models: Challenges and opportunities for geographical and environmental education. *International Research in Geographical and Environmental Education*, 32(2), 85-89. <https://doi.org/10.1080/10382046.2023.2194036>
- Chounta, I., Bardone, E., Raudsep, A., & Pedaste, M. (2021). Exploring teachers' perceptions of artificial intelligence as a tool to support their practice in Estonian K-12 education. *International Journal of Artificial Intelligence in Education*, 32(3), 725-755. <https://doi.org/10.1007/s40593-021-00243-5>
- Cooper, G. (2023). Examining science education in ChatGPT: An exploratory study of generative artificial intelligence. *Journal of Science Education and Technology*, 32(3), 444-452. <https://doi.org/10.1007/s10956-023-10039-y>
- Dilzhan, B. (2024). Teaching English and artificial intelligence: EFL teachers' perceptions and use of ChatGPT (Dissertation). SDU University.
- Dwivedi, Y. K., Pandey, N., Currie, W., & Micu, A. (2023). Leveraging ChatGPT and other generative artificial intelligence (AI)-based applications in the hospitality and tourism industry: Practices, challenges and research agenda. *International Journal of Contemporary Hospitality Management*, 36(1), 1-12. <https://doi.org/10.1108/ijchm-05-2023-0686>
- Han, A., Zhou, X., Cai, Z., Han, S., Ko, R., Corrigan, S., & Pepler, K. A. (2024). Teachers, parents, and students' perspectives on integrating generative AI into elementary literacy education. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (pp. 1-17).
- Hıncz, G. (2024). A year of generative AI in English language teaching and learning: A case study. *Journal of Research on Technology in Education*, 1-21. <https://doi.org/10.1080/15391523.2024.2404132>
- Holmes, W., & Miao, F. (2023). Guidance for generative AI in education and research. UNESCO Publishing.
- Holmes, W., Bialik, M., & Fadel, C. (2021). Artificial intelligence in education: Promises and implications for teaching and learning. Center for Curriculum Redesign.
- Hwang, G. J., & Chen, N. S. (2023). Exploring the potential of generative artificial intelligence in education: Applications, challenges, and future research directions. *Journal of Educational Technology & Society*, 26(2).

Jauhiainen, J. S., & Guerra, A. G. (2023). Generative AI and ChatGPT in school children's education: Evidence from a school lesson. *Sustainability*, 15(18), 14025. <https://doi.org/10.3390/su151814025>

Kitonga, M. B., Nageh, H. M., Nduta, B. G., Mafuwane, B. M., & Mochekele, M. (2024). Education and professional development. *Cari Journals USA LLC*.

Kong, S. C., Yang, Y., & Hou, C. (2024). Examining teachers' behavioural intention of using generative artificial intelligence tools for teaching and learning based on the extended technology acceptance model. *Computers and Education: Artificial Intelligence*, 7, 100328. <https://doi.org/10.1016/j.caeai.2024.100328>

Labajová, L. (2023). The state of AI: Exploring the perceptions, credibility, and trustworthiness of the users towards AI-generated content.

Le-Nguyen, H. T., & Tran, T. T. (2024). Generative AI in terms of its ethical problems for both teachers and learners: Striking a balance. In *generative AI in teaching and learning* (pp. 144-173). IGI Global.

Liao, H., Xiao, H., & Hu, B. (2023). Revolutionizing ESL teaching with generative artificial intelligence: Take ChatGPT as an example. *International Journal of New Developments in Education*, 5(20), 39-46. <https://doi.org/10.25236/ijnde.2023.052008>

Liu, J. (2024). Enhancing English language education through big data analytics and generative AI. *Journal of Web Engineering*, 227-250. <https://doi.org/10.13052/jwe1540-9589.2322>

Liu, M., Ren, Y., Nyagoga, L. M., Stonier, F., Wu, Z., & Yu, L. (2023). Future of education in the era of generative artificial intelligence: Consensus among Chinese scholars on applications of ChatGPT in schools. *Future in Educational Research*, 1(1), 72-101. <https://doi.org/10.1002/fer3.10>

Luckin, R. (2017). *Machine learning and human intelligence: The future of education for the 21st century*. UCL Press.

Lv, Z. (2023). Generative artificial intelligence in the metaverse era. *Cognitive Robotics*, 3, 208-217. <https://doi.org/10.1016/j.cogr.2023.06.001>

Mannuru, N. R., Shahriar, S., Teel, Z. A., Wang, T., Lund, B. D., Tijani, S., Pohboon, C. O., Agbaji, D., Alhassan, J., Galley, J., Kousari, R., Ogbadu-Oladapo, L., Saurav, S. K., Srivastava, A., Tummuru, S. P., Uppala, S., & Vaidya, P. (2023). Artificial intelligence in developing countries: The impact of generative artificial intelligence (AI) technologies for development. *Information Development*. <https://doi.org/10.1177/02666669231200628>

Mao, J., Chen, B., & Liu, J. C. (2023). Generative artificial intelligence in education and its implications for assessment. *TechTrends*, 68(1), 58-66. <https://doi.org/10.1007/s11528-023-00911-4>

- McIntosh, T. R., Susnjak, T., Liu, T., Watters, P., Xu, D., Liu, D., & Halgamuge, M. N. (2025). From Google Gemini to OpenAI Q star: A survey on reshaping the generative artificial intelligence research landscape. *Technologies*, 13(2), 51. <https://doi.org/10.3390/technologies13020051>
- Mittal, U., Sai, S., Chamola, V., & Sangwan, D. (2024). A comprehensive review on generative AI for education. *IEEE Access*, 12, 142733-142759. <https://doi.org/10.1109/access.2024.3468368>
- Moorhouse, B. L., & Kohnke, L. (2024). The effects of generative AI on initial language teacher education: The perceptions of teacher educators. *System*, 122, 103290. <https://doi.org/10.1016/j.system.2024.103290>
- Nedungadi, P., Tang, K., & Raman, R. (2024). The transformative power of generative artificial intelligence for achieving the sustainable development goal of quality education. *Sustainability*, 16(22), 9779. <https://doi.org/10.3390/su16229779>
- Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, 381(6654), 187-192. <https://doi.org/10.1126/science.adh2586>
- Ooi, K., Tan, G. W., Al-Emran, M., Al-Sharafi, M. A., Capatina, A., Chakraborty, A., Dwivedi, Y. K., Huang, T., Kar, A. K., Lee, V., Loh, X., Micu, A., Mikalef, P., Mogaji, E., Pandey, N., Raman, R., Rana, N. P., Sarker, P., Sharma, A., & Wong, L. (2023). The potential of generative artificial intelligence across disciplines: Perspectives and future directions. *Journal of Computer Information Systems*, 65(1), 76-107. <https://doi.org/10.1080/08874417.2023.2261010>
- Pavlik, J. V. (2023). Collaborating with ChatGPT: Considering the implications of generative artificial intelligence for journalism and media education. *Journalism and Mass Communication Educator*, 78(1), 84-93. <https://doi.org/10.1177/10776958221149577>
- Qadir, J. (2023). Engineering education in the era of ChatGPT: Promise and pitfalls of generative AI for education. 2023 IEEE Global Engineering Education Conference (EDUCON), 1-9. <https://doi.org/10.1109/educon54358.2023.10125121>
- Samala, A. D., Rawas, S., Wang, T., Reed, J. M., Kim, J., Howard, N., & Ertz, M. (2024). Unveiling the landscape of generative artificial intelligence in education: A comprehensive taxonomy of applications, challenges, and future prospects. *Education and Information Technologies*, 30(3), 3239-3278. <https://doi.org/10.1007/s10639-024-12936-0>
- Schmidt, T., & Strassner, T. (2022). Artificial intelligence in foreign language learning and teaching. *Anglistik*, 33(1), 165-184. <https://doi.org/10.33675/angl/2022/1/14>
- Sengar, S. S., Hasan, A. B., Kumar, S., & Carroll, F. (2024). Generative artificial intelligence: A systematic review and applications. *Multimedia Tools and Applications*. <https://doi.org/10.1007/s11042-024-20016-1>
- Song, Y., Kim, J., Xing, W., Liu, Z., Li, C., & Oh, H. (2025). Elementary school students' and teachers' perceptions toward creative mathematical writing with generative AI. *Journal of Research on Technology in Education*, 1-23. <https://doi.org/10.1080/15391523.2025.2455057>

Su, J., & Yang, W. (2023). Unlocking the power of ChatGPT: A framework for applying generative AI in education. *ECNU Review of Education*, 6(3), 355-366. <https://doi.org/10.1177/20965311231168423>

Tai, T., & Chen, H. H. (2024). Improving elementary EFL speaking skills with generative AI chatbots: Exploring individual and paired interactions. *Computers and Education*, 220, 105112. <https://doi.org/10.1016/j.compedu.2024.105112>

Tseng, J., Chai, C. S., Tan, L., & Park, M. (2020). A critical review of research on technological pedagogical content knowledge (TPACK) in language teaching. *Computer Assisted Language Learning*, 35(4), 948-971. <https://doi.org/10.1080/09588221.2020.1868531>

Verma, G., Campbell, T., Melville, W., & Park, B. (2023). Navigating opportunities and challenges of artificial intelligence: ChatGPT and generative models in science teacher education. *Journal of Science Teacher Education*, 34(8), 793-798. <https://doi.org/10.1080/1046560x.2023.2263251>

Walczak, K., & Cellary, W. (2023). Challenges for higher education in the era of widespread access to generative AI. *Economics and Business Review*, 9(2). <https://doi.org/10.18559/eb.2023.2.743>

Wu, Y. (2023). Integrating generative AI in education: How ChatGPT brings challenges for future learning and teaching. *Journal of Advanced Research in Education*, 2(4), 6-10. <https://doi.org/10.56397/jare.2023.07.02>

Yilmaz, R., & Karaoglan Yilmaz, F. G. (2023). The effect of generative artificial intelligence (AI)-based tool use on students' computational thinking skills, programming self-efficacy and motivation. *Computers and Education: Artificial Intelligence*, 4, 100147. <https://doi.org/10.1016/j.caeai.2023.100147>

Yu, H., & Guo, Y. (2023). Generative artificial intelligence empowers educational reform: Current status, issues, and prospects. *Frontiers in Education*, 8. <https://doi.org/10.3389/feduc.2023.1183162>