LEVERAGING GENERATIVE AI IN ELT: TEACHERS' INTEGRATION STRATEGIES AND PEDAGOGICAL ADAPTATIONS

Finita Dewi
Faculty of Language and Literature Education, Universitas Pendidikan Indonesia, Indonesia
Corresponding Author Email: finita@upi.edu

Article Info	Abstract
	This study investigates how English teachers integrate Generative AI—ChatGPT, Diffit, and Brisk—into their instructional practices, focusing on their strategies, challenges and pedagogical adaptations. These tools were selected based on their distinctive functionalities: ChatGPT for content generation, Diffit for text simplification, and Brisk for creating multimedia presentations and teaching aids. Using a qualitative multiple case study approach, data were collected from twelve primary and junior high school ELT teachers, setting through reflective journals, focus group discussions, and document analysis. The findings of this study showed that although GAI offered strengths such as fostering creativity, enhancing accessibility, and engaging students, the tools also have some limitations in terms of content inaccuracy, oversimplification, and lack of contextual adaptability. Therefore, teacher intervention is required in its implementation. In this study, teachers used reflection-in-action to make real-time adjustments, reflection-on-action to refine the integration strategies, and collaborative reflection to exchange best practices. Additionally, findings highlight teacher agency's importance in leveraging AI for lesson planning and differentiated instruction. It further highlights the need for professional development in AI literacy to equip educators with skills to evaluate and adapt AI-generated content critically. This research contributes to the growing discourse on AI in Education, demonstrating how generative AI can function as a teaching aid rather than a replacement for pedagogical experts.

INTRODUCTION

The rapid advancement of generative artificial intelligence (GAI) has ushered into a new era of educational innovation, transforming traditional teaching and learning processes. GAI enables educators to automate repetitive tasks, design customized materials, and foster creativity in classroom practices. In the field of English Language Teaching (ELT), AI-powered tools have gained increasing attention for their potential to enhance lesson planning, develop instructional material, and provide personalized learning experiences (Mishra et al., 2023; Zawacki-Richter et al., 2019). Teachers frequently contend with diverse learner needs and resource-intensive preparation. GAI has shown immense potential to streamline workflows and enhance instructional quality (Cogo et al., 2024; Susanto et al., 2024). These tools offer innovative ways for teachers to generate learning materials, adapt texts for different proficiency levels, and create engaging multimedia content.

The integration of AI technologies in Education has profound implications for the role of teachers. While AI can handle tasks such as grading and generating instructional content, it does not replace the nuanced, human-centred aspects of teaching. As highlighted by (Krushinskaia et al., 2023), AI's primary role should be to augment rather than substitute teachers' expertise, empowering them to focus on higher-order skills such as mentoring, facilitating discussions, and fostering critical thinking. However, these benefits are

accompanied by some challenges. The effective integration of GAI in ELT classrooms requires teacher agency, reflective adaptation, and ongoing pedagogical decision-making (Nyaaba & Zhaı, 2024). Teachers often lack the training to prompt these tools effectively and critically assess the quality of AI-generated outputs.

To thrive in this evolving landscape, teachers must develop competencies in navigating AI, including data literacy and pedagogical adaptation. The reflective practice serves as a cornerstone for this transformation, offering teachers a framework to assess what works, identify challenges, and implement iterative improvements in their integration of AI (Mishra et al., 2023; Nyaaba & Zhaı, 2024). This reflective process is particularly relevant in ELT, where the effectiveness of generative AI hinges on its ability to align with language teaching objectives, support student engagement, and maintain cultural sensitivity.

While research on AI in Education has rapidly expanded, much of the focus emphasizes student engagement and AI-driven assessment, with fewer investigations on how much teachers actively engage with AI to enhance their instructional strategies (Celik, 2023; Ghimire et al., 2024). Recent studies highlight that AI-generated content often requires modification, as the output may lack contextual depth, cultural sensitivity, or pedagogical appropriateness (Cogo et al., 2024). However, little has been done to explore how teachers reflect on their adaptation to AI-generated materials to meet their classroom needs.

The intersection of AI and language learning requires a technological pedagogical content knowledge (TPACK) framework, where teachers must integrate not only technological knowledge (TK) but also technological content knowledge (TCK) and technological pedagogical knowledge (TPK) to ensure AI is effectively used in language instruction (Albion et al., 2010; Dewi, 2019; Tai, 2015). Recent studies have emphasized that teachers who effectively integrate AI into ELT do not simply use technology as a content generator but as a pedagogical enabler that enhances instructional quality (Dewi & Hamied, 2024; Novawan et al., 2024). AI-generated materials must be critically adopted to avoid failure aligning with language acquisition principles, which might lead to pedagogical challenges.

In addition, there is an urgent need to pay attention to the ethical aspects, including bias in AI-generated content, the accuracy of AI-generated information, and data privacy (Brandão et al., 2024). GAI may produce incorrect facts in its responses, reflect biases from the training data, or even fail to recognise regional linguistic nuances. Moreover, the text simplification process can unintentionally remove essential content or change meaning, affecting the content's comprehension and cultural relevance. The concerns show the need for teacher oversight and AI literacy training, ensuring the responsible utilization of AI in the teaching context (Nyaaba & Zhaı, 2024).

Despite these challenges, there has been no extensive exploration of how teachers navigate AI limitations through reflective practices. While many studies have addressed AI's technical and ethical challenges, there remains a gap in understanding how teachers critically evaluate, modify, and adapt AI-generated content, especially in the Indonesian context (Williyan et al., 2024). In borrowing (Schon, 1983) terms, teachers often engaged in reflection-in-action (adjusting AI-generated materials during lessons), and reflection-on-action (revising strategies post-lesson), but these processes remain understudied. Understanding how teachers apply these reflective strategies can provide insights into best practices for AI integration in ELT.

Furthermore, professional development's role in teachers' AI literacy is still underexamine. Studies suggest that teachers who receive training on AI and prompt engineering are better equipped to maximize AI's benefits and mitigate the risks (Tunjera & Chigona, 2023). However, ELT teachers do not have enough training in AI literacy; they rely on trial-and-error approaches to adapt AI to their teaching (Susanto et al., 2024). Thus, by examining how

teachers learn, experiment, and share strategies for integrating AI, this study could contribute to the growing discourse of teacher agency in AI-enhanced ELT.

To address these gaps, this study examines how ELT teachers integrate GAI into their teaching practices—the challenges they encounter and the reflective strategies they employ to adapt AI-generated content. By using the Technological Pedagogical Content Knowledge (TPACK) framework (Mishra & Koehler, 2006) and Reflective Practice Theory (Schon, 1983), this study seeks to answer three research questions: 1) *How do ELT teachers integrate GAI into their lessons?, and in what ways does this integration reflect their TPACK?,* 2) *What are the perceived strengths and Weaknesses of GAI in supporting ELT?,* 3) *How do reflective practices help teachers adapt and overcome challenges when using GAI?*

By combining TPACK with reflective practice, this study offers a dual-lens perspective that bridges theoretical and practical domains. It provides actionable insights for educators, policymakers, and AI developers, emphasizing the critical role of teachers in ensuring that generative AI not only augments educational outcomes but also upholds the core values of effective pedagogy. This research contributes to a growing body of literature that positions teachers as central to the successful integration of AI in Education, highlighting the importance of empowering them with the tools, training, and reflective frameworks needed to navigate this rapidly evolving landscape.

RESEARCH METHOD
Research Design

This study employed a qualitative case study approach (Yin, 2009) to investigate how ELT teachers integrate GAI into their instructional practices. A case study design was chosen because it allows for capturing in-depth insights into teachers' experiences with these tools in diverse classroom settings. It also enables an understanding of how teachers navigate the affordances and constraints of GAI within their specific pedagogical environments and provides flexibility to capture diverse teacher perspectives and adaptations across different contexts (Creswell & Creswell, 2017; Stake, 2010).

This study was conducted within a technology-focused Teacher Professional Development (TPD) program organized by a Bandung, West Java Province, Indonesia school foundation. This school foundation manages multiple schools across primary, junior high, and senior high school levels. The TPD program was aimed at updating teachers' competence with some emphasis on exploring the potential of GAI as a teaching assistant. Teachers participating in this TPD were introduced to several GAIs, explored the potentials of the GAI through hands-on experience, and were encouraged to integrate them into their teaching practices. This study occurred within the authentic professional learning environment, allowing researchers to examine how teachers engaged with AI, reflected on its utilization, and adapted their instructional approaches.

This study is framed by the TPACK framework (Mishra & Koehler, 2006) and guided the analysis of how technological, pedagogical, and content knowledge influenced the integration strategies. Additionally, Schön's (1983) reflective practice theory informs the analysis of how teachers evaluate, modify, and refine AI-generated content through reflection-in-action and reflection-on-action.

Participants

The study involved twelve English Language Teaching (ELT) teachers participating in the technology-focused TPD program. These teachers were selected through purposive sampling (Patton, 2002) to ensure representation across teaching and GAI proficiency levels. Their participation in the TPD program meant they had been exposed to GAI and were actively encouraged to experiment with GAI in their lesson planning and classroom activities. To

capture the participants' GAI integration experiences, they were categorized based on their self-reported familiarity with GAI. There are three categories of participants: (a) Beginner (4 participants): Teachers with minimal prior exposure to GAI engaging with GAI for the first time during the TPD program; (b) Intermediate (5 participants): Teachers with some prior experience using GAI in lesson planning but still exploring its full potential; and (c) Advanced (3 participants): Teachers who had previously integrated GAI into their teaching were comfortable critically customizing AI-generated content.

Selection of Generative Artificial Intelligence

The teachers were introduced to three different GAI, ChatGPT, Brisks and Diffit, as part of the TPD program. These tools were selected for their potential to provide different affordances aligning with different instructional needs.

ChatGPT (https://chatgpt.com/) was chosen because of its capabilities to generate content required by teachers to support the teaching-learning process. ChatGPT can help teachers design lesson plans, create vocabulary and grammar exercises, write prompts, provide sample dialogues, and simplify and expand texts to suit students' language proficiency levels.

Diffit (https://web.diffit.me/) was chosen due to its ability to adjust reading materials based on student proficiency. This tool also allows teachers to modify text complexity while maintaining meaning and coherence, making it particularly useful for mixed-ability classrooms. It is also possible for teachers to generate text from any given topic, a list of key vocabulary words, and YouTube videos. The third GAI tool is Brisk. It was selected for its potential to enhance lesson delivery through visually engaging and structured content.

Brisk (https://www.briskteaching.com/) allows teachers to create interactive lesson slides, organize instructional materials, and create quizzes, lesson plans, questions prompts, and feedback for students' writing pieces. The selection of these three GAI tools allowed the researcher to comprehensively analyze AI integration in ELT, covering creation, differentiation, and engagement from the teaching and learning process.

Data Collection

This study employed three complementary data collection methods—reflective journals, focus group discussions (FGDs), and document analysis—to achieve robust data triangulation and a comprehensive understanding of teachers' experiences with AI integration in English Language Teaching (ELT). Each instrument was carefully selected to capture distinct yet interconnected dimensions of teacher practice, reflection, and resource use. Reflective journals provided insights into individual teachers' thought processes and adaptive strategies, while FGDs enabled collaborative reflection and the exchange of ideas. Document analysis, on the other hand, offered contextual data on instructional materials and AI-generated outputs. To ensure the validity and reliability of the instruments, a preliminary phase was conducted involving three ELT teachers outside the main participant group. These pilot participants reviewed the journal prompts, FGD questions, and document analysis checklist, offering feedback on clarity, relevance, and coherence with the study's objectives. Based on their input, revisions were made to enhance the quality and alignment of each instrument with the research goals.

Reflective journals

Reflective journals captured the teachers' detailed experiences with GAI by eliciting their critical self-assessment and iterative learning (Richards & Farrell, 2011). Teachers were asked to write their weekly reflective journals for four weeks to record their experiences integrating GAI into their lessons and classroom instruction. The reflective prompts (See Table 1) were designed using the TPACK Framework.

Table 1
Reflective Journal Prompts

Reflection Phase	Journal Prompt	TPACK Dimension
Pre-lesson Planning	What were your primary lesson objectives, and how did you plan to use GAI?	TK-TCK
	How did you assess whether the AI-generated materials were appropriate for your students?	TCK
During the Lesson	How did students respond to the AI-generated materials? Did the tool support student engagement and understanding?	TPK
	Did you need to make any real-time adjustments to the AI-generated materials? If so, why?	Reflection-in-action (TPK-TPACK)
Post-lesson Reflection	How did GAI contribute to your lesson's success? What were its limitations?	TPACK
	What improvements would you make in your next lesson?	Reflection-on-Action (TPACK Development)

Focus Group Discussions (FGDs)

FGDs are commonly recognized methods for capturing group dynamics and collective problem-solving (Morgan, 1997). FGDs were conducted after the implementation phase to provide a collaborative space for teachers to share and compare their GAI integration experiences. Three semi-structured FGDs were organized, one for each school level (primary, junior high, senior high), allowing the participants to share and discuss challenges and pedagogical strategies during the implementation phase. The discussion questions (see Table 2) during FGDs were aligned with the information gathered from the teachers' reflective journals.

Table 2
FGD Question Prompts

Theme	FGD Question	TPACK Dimension
GAI usability	What was your overall experience using ChatGPT, Diffit, or Brisk in your teaching?	TK
Content Accuracy	Have you encountered inaccuracies or overly simplified AI-generated content? How did you address them?	TCK
Instructional Integration	In what ways did GAI enhance or hinder your instructional strategies	TPK
Challenges and Adjustments	How did you adapt AI-generated materials to fit your lesson objectives better?	TPACK
Reflective Practice	What strategies helped you refine your GAI usage over time?	Reflection-on-Action (TPACK Development)

Document Analysis

Document analysis, as defined by Bowen (2009), is a systematic technique for examining written, visual, or digital artifacts to evaluate their relevance and alignment with specific objectives—in this case, educational goals. In this study, participants submitted various AI-generated teaching materials, including lesson plans, student worksheets, and multimedia presentation slides. These materials were analyzed using a structured evaluation checklist (refer to Table 3), designed to assess the overall quality, practicality, and pedagogical coherence of the content. The analysis aimed to determine how effectively AI tools supported instructional design and whether the outputs aligned with curriculum standards and sound teaching practices.

Table 3
Document Analysis Evaluation Criteria

Evaluation Criteria	TPACK Dimension
Creativity, functionality, and minimal editing needs	TK
Alignment with lesson objectives and language proficiency levels	TCK
Clarity, usability, and engagement potential	TPK
Customization and Adaptability	TPACK
Accuracy and Ethical Consideration	TCK & TPACK

Data Analysis

Thematic analysis was employed to analyze data from reflective journals and FGDs, following the six-step approach proposed by (Braun & Clarke, 2006). This method facilitated the identification of patterns and themes across the TPACK dimensions (TK, PK, CK) and reflective practices (Reflection-in-action, Reflection-on-action). Coding categories were derived deductively from the frameworks and inductively from participants' narratives, ensuring a comprehensive analysis. The analysis was conducted manually to allow greater researcher immersion in the data and a more nuanced interpretation of emerging patterns. Each transcript generated from FGDs and reflective journal entries was read multiple times to ensure the researcher's familiarity with the data. Coding was performed manually by highlighting key phrases and categorizing excerpts into thematic groups based on the alignment with TPACK dimensions. Codes were then cross-referenced to ensure consistency, and adjustments were made to refine theme categorization.

Document analysis used the structured checklist to score materials on a five-point scale, with comments providing qualitative insights. Scores were synthesized with qualitative findings to compare the strengths and weaknesses of each tool across different teaching contexts.

Ethical Considerations The study adhered to ethical research practices, as Creswell and Creswell (2018) outlined. Informed consent was obtained from all participants, and their identities were anonymized to protect confidentiality. Teachers were informed of their right to withdraw at any stage without penalty.

Reflective journals allowed for introspective, detailed narratives of teachers' experiences, while FGDs provided a platform for collective reflection and triangulation of insights. Document analysis added an objective layer of evaluation, capturing how AI-generated materials aligned with pedagogical goals. This combination of methods, rooted in established qualitative research practices, ensured a robust and comprehensive exploration of the integration of GAI in ELT.

RESEARCH FINDINGS AND DISCUSSION
Findings

The findings section showcases the results of the studies, which attempted to address the three research questions through data generated from reflective journals, FGDs, and document analysis. The findings of this study show the extent to which teachers integrated GAI into their teaching, their perception towards the strengths and weaknesses of ChatGPT, Diffit, and Brisk, and the essential role of reflective practice in navigating the challenges. This result highlights the importance of teacher agency and adaptive strategies in leveraging the potential of GAI for English language teaching.

Findings 1: Teachers' Integration of Generative AI Tools into ELT

This study revealed that teachers have different ways of integrating different GAI into their teaching and learning process in the classroom. The way they integrated them was guided by their pre-defined instructional objectives and classroom practices. Data collected from reflective journals (RJ), focus group discussions (FGDs), and document analysis (DA) shows the unique capabilities and limitations of each tool in the integration patterns.

ChatGPT: A Flexible Content Generator

Among the three tools, ChatGPT was the most flexible tool, especially for creating lesson plans, writing prompts, and idea generation. The teacher used ChatGPT to promote group discussions, argumentative writing, and a project requiring higher-order thinking and creativity. Using ChatGPT, one of the participants (RJ4) explained how to generate debate topics related to environmental preservation:

> *"ChatGPT allowed me to introduce multiple perspectives quickly, which helped students think critically about both sides of an argument."*

Similarly, in an FGD, a teacher (FGD2-P3) explained:

> *"I used ChatGPT to generate brainstorming ideas for essay writing. It saved me time, but I had to adjust the output because some examples were too advanced for my Grade 9 students."*

The document analysis (DA-C1) confirmed that although ChatGPT outputs were rich in content, editing was often needed to match class objectives and student competency levels. The teachers mentioned that ChatGPT was utilized for writing prompts tailored to students' interests, modelling answers to help students understand complex writing structures, and generating diverse viewpoints for classroom debates. However, its effectiveness depended heavily on how teachers crafted prompts, as outputs varied in quality and relevance.

Diffit: Simplifying Texts for Accessibility

Diffit was widely used to create simplified reading passages and comprehension tasks. Teachers used the application of the tool to adapt texts for beginner and intermediate learners, ensuring accessibility without compromising core content. A reflective journal entry (RJ2) described how a teacher simplified a science article for a Grade 5 class:

> *"The original passage on the water cycle was too technical for my students, but Diffit's output made it more accessible while retaining the main ideas."*

In an FGD, another teacher (FGD1-P5) emphasized Diffit's role in reducing preparation time:

> *"Diffit simplified a historical text in minutes, saving me hours of manual adaptation."*

However, document analysis (DA-D4) showed several issues with oversimplification, in which the teacher found some critical contextual details were omitted, affecting the richness of the text. Teachers mainly utilized Diffit to simplify complex reading materials for differentiated instruction and create comprehension questions tailored to different students' proficiency levels. Despite the strengths of this tool, some participants highlighted that the autogenerated questions were considered too essential and required supplementary editing for higher-order thinking activities.

Brisk: Enhancing Visual Engagement

Brisk was primarily used to create multimedia presentations to introduce or reinforce concepts. Its visual appeal was particularly valued in lessons that involved abstract or challenging topics. A teacher (RJ6) highlighted Brisk's effectiveness in a Grade 12 literature class:

"The slides Brisk created for 'The Ugly Duckling" story helped my students
visualize the symbolism, making abstract themes more accessible."

Another participant (FGD3-P2) shared:

"Brisk saved me time when preparing a grammar presentation. The visuals kept my
students engaged, though I had to add more details to make the content substantial."

Document analysis (DA-B2) showed that while Brisk presentations were polished and visually engaging, they often relied on generic templates that limited customization for advanced topics. Teachers used Brisk to create visually appealing slides for lessons on literary themes, grammar rules, or historical contexts and support discussion-based activities by presenting key points in an engaging format. However, its reliance on pre-designed templates meant additional teacher input was often needed to achieve sufficient depth.

The integration of ChatGPT, Diffit, and Brisk demonstrated that teachers employed these tools strategically, matching their unique affordances to specific classroom needs. ChatGPT was most effective for content generation in creative and critical thinking tasks. Meanwhile, Diffit facilitated accessibility, especially for younger or less proficient learners. Brisk excelled in visual engagement but required teacher intervention to meet advanced learning objectives. Across all tools, reflective practices played a crucial role in optimizing their use. Teachers frequently adjusted outputs to align with their students' proficiency levels and lesson goals, highlighting the importance of teacher expertise in integrating generative AI into ELT.

Findings 2: Perceived Strengths and Weaknesses of the Tools

The study revealed distinct strengths and weaknesses of ChatGPT, Diffit, and Brisk in supporting English Language Teaching (ELT). These perceptions emerged from reflective journals (RJ), focus group discussions (FGDs), and document analysis (DA), highlighting the tools' contributions and limitations in addressing lesson objectives, enhancing student engagement, and supporting instructional strategies.

Strengths of the Tools

ChatGPT was recognized for its ability to produce diverse teaching materials, ranging from essay prompts to brainstorming activities. Its capacity to generate content aligned with various teaching objectives made it particularly effective for higher-level ELT tasks. In an FGD, one teacher (FGD2-P3) shared:

"ChatGPT helped me provide multiple examples of argumentative writing structures,
encouraging my students to think critically about different perspectives."

Reflective journals also highlighted its creative potential, with a participant (RJ5) noting:

"The tool's flexibility allowed me to tailor writing tasks to my students' interests, making the
activity more engaging."

Document analysis (DA-C1) showed that ChatGPT outputs were rich in content and varied in scope, making them suitable for diverse classroom contexts. Diffit's ability to simplify texts was frequently praised for making advanced content accessible to lower-proficiency learners. Teachers emphasized its role in supporting differentiated instruction. A teacher (RJ3) reflected:

"Diffit simplified a challenging passage on climate change into language my Grade 7 students could understand without losing the key ideas."

Brisk was highly valued for its ability to create visually appealing presentations that enhanced student engagement. Teachers noted that it was particularly effective for introducing abstract or thematic content. A teacher (FGD3-P1) described:

"Brisk helped me explain the abstract concept with visuals that my students found relatable."

Weaknesses of the Tools

While ChatGPT was praised for its flexibility, its outputs were often inconsistent in quality and required substantial editing to ensure alignment with lesson objectives and student proficiency levels. One teacher (FGD2-P2) explained:

"Sometimes the examples ChatGPT generates are too advanced, and I spend more time simplifying them than I would be writing them myself."

A journal entry (RJ4) reflected similar concerns:

"The tool occasionally produces inaccurate or irrelevant content, which disrupts my lesson planning process."

Document analysis (DA-C3) confirmed these challenges, with several outputs requiring extensive teacher mediation to ensure clarity and appropriateness.

Diffit was frequently critiqued for oversimplifying texts, resulting in a loss of nuance or contextual richness. Additionally, its auto-generated comprehension questions were often repetitive or lacked higher-order thinking elements. A teacher (FGD1-P3) shared:

"While Diffit simplifies texts effectively, it sometimes strips them of critical details essential for deeper understanding."

The findings suggest that while each tool offered unique advantages, their effectiveness was contingent on teacher mediation to address their limitations. The tools' perceived strengths and weaknesses are summarized below:

Table 4
Summary of Teachers' Evaluations of ChatGPT, Diffit, and Brisk

Tool	Strengths	Weaknesses	Perceived Usefulness
ChatGPT	Generate varied ELT materials and support multiple language proficiency levels.	It requires careful, prompt engineering; the content is sometimes too complex or off-topic.	Helpful in creating initial drafts of exercises, reading passages, and discussion prompts
Diffit	Simplifies texts for different proficiency levels, easy to use	Can oversimplify content, losing important details or context	Best for reading material adaptation and differentiated instructions
Brisk	Creates structured presentations with AI-generated slides	Slides can be generic or not ELT-specific, requiring customization	It is helpful for lesson delivery but requires teacher modifications

Findings 3: The Role of Reflective Practices in Adapting to Challenges

Reflective practices emerged as a central strategy for teachers to address the limitations of ChatGPT, Diffit, and Brisk. Data from reflective journals (RJ), focus group discussions (FGDs), and document analysis (DA) demonstrated how teachers engaged in both reflection-in-action (real-time problem-solving during lessons) and reflection-on-action (post-lesson evaluations and planning). Collaborative reflections through FGDs further highlighted shared strategies and collective problem-solving among participants.

Reflection-in-Action: Real-Time Adjustments

Teachers frequently identified limitations in AI-generated materials during lessons and made on-the-spot modifications to maintain lesson flow and achieve objectives. One teacher (RJ4) described adapting ChatGPT outputs during a lesson:

> *"Some of the examples were too advanced for my students. I paused and simplified the sentences verbally to make them more accessible."*

Similarly, in an FGD, a participant (FGD2-P1) shared their experience with Diffit:
> *"When students struggled with the oversimplified text, I used the original passage to explain some missing details, balancing both versions during the lesson."*

Document analysis (DA-D3) supported these observations, revealing teacher annotations and edits made to Diffit outputs before or during lessons to fill gaps or clarify ambiguous content.

Reflection-on-Action: Post-Lesson Evaluations

After lessons, teachers engaged in reflective evaluations to analyze what worked, identify challenges, and plan improvements for future implementations. A reflective journal entry (RJ6) detailed a teacher's post-lesson analysis of Brisk-generated slides:

> *"The visuals were engaging, but the lack of depth in the content reduced their impact. Next time, I'll add detailed textual explanations to the slides in advance."*

Another participant (FGD3-P3) explained how they refined their use of ChatGPT:

> *"After realizing my initial prompts were too broad, I used specific keywords and examples to guide the tool's output effectively."*

Document analysis (DA-C2) corroborated this reflective process, showing improved alignment between initial and revised ChatGPT outputs, with teachers making edits to better suit learning objectives.

Collaborative Reflection: Shared Strategies and Problem-Solving

FGDs revealed that teachers valued sharing experiences and solutions with colleagues, fostering a collaborative learning environment. Many participants highlighted how peer discussions helped refine their approaches to using the tools. One participant (FGD1-P4) shared a technique for improving ChatGPT outputs:

> *"Using scenario-based prompts helped me get more relevant content. For instance, asking for a dialogue about ordering food at a restaurant produced a realistic conversation for my Grade 7 class."*

Another teacher (FGD3-P1) discussed their collaborative approach to enhancing Diffit outputs:

"I combined the simplified text with comprehension questions my colleague created manually. Together, we made a resource that worked for all levels."

Reflective journals (RJ3, RJ5) also noted the influence of collaborative discussions, with participants often implementing suggestions shared during FGDs to improve their use of generative AI tools.

Teachers' engagement in reflective practices—individually and collaboratively—was instrumental in addressing the challenges posed by generative AI tools. These practices enabled iterative improvements, ensuring AI outputs aligned more closely with pedagogical goals. The reflective process can be summarized as follows:

Table 5
Summary of Reflective Process

Type of Reflection	Description	Example
Reflection-in-Action	Real-time problem-solving during lessons to address gaps or inaccuracies in AI-generated materials.	Simplifying ChatGPT examples on the spot to match student proficiency (RJ4).
Reflection-on-Action	Post-lesson evaluations and planning for future improvements in AI tool integration.	Revising Brisk slides to add textual explanations after observing limited content depth during lessons (RJ6).
Collaborative Reflection	Peer discussions and shared strategies for optimizing tool usage.	Combining Diffit outputs with manually created comprehension questions, shared during an FGD, to support differentiated instruction (FGD1-P4).

Discussion

The findings of this study provide critical insights into how generative AI tools—ChatGPT, Diffit, and Brisk—are integrated into English Language Teaching (ELT), the strengths and weaknesses of these tools, and the role of reflective practices in overcoming challenges. This section discusses these findings in the context of relevant theories, such as Technological Pedagogical Content Knowledge (TPACK) and Reflective Practice, as well as insights from recent literature on AI in Education.

RQ1: *How do ELT teachers integrate GAI into their lessons? In what ways does this integration reflect their TPACK?*

The study found that teachers integrated ChatGPT, Diffit, and Brisk into their teaching in ways that reflected their unique affordances. ChatGPT's flexibility made it ideal for generating diverse materials, Diffit enhanced accessibility by simplifying complex texts, and Brisk engaged students through visual aids. However, the need for substantial teacher intervention across all tools aligns with the TPACK framework, which highlights the critical interplay of technological knowledge (TK), pedagogical knowledge (PK), and content knowledge (CK).

Teachers demonstrated strong TK navigating the AI tools' interfaces, but their success depended on their PK and CK tailoring outputs to specific classroom needs. For instance, while ChatGPT excelled at producing varied content, its reliance on precise prompts and inconsistent outputs required teachers to draw on their PK to scaffold learning effectively. Similarly, Diffit's ability to simplify texts depended on teachers' CK to supplement lost details and maintain depth, particularly in higher-level contexts. Brisk's visual appeal supported PK by engaging

students in discussion-based activities, but its limited customization options constrained its CK alignment.

These findings echo the work of Mishra et al. (2023), who emphasized that integrating AI tools requires a nuanced understanding of the intersections between technology, pedagogy, and content. Similarly, Susanto et al. (2024) observed that AI tools in ELT often serve as a starting point rather than a complete solution, requiring teachers to adapt outputs to meet educational objectives critically. These results are consistent with (Dewi, 2019), who found that Indonesian EFL teachers needed to modify technology-generated content to align with ELT principles. Similarly, Cogo et al. (2024) highlight that AI-generated lesson materials often lack pedagogical depth, requiring teachers to mediate and adapt content before classroom implementation.

RQ2: *What are GAI's perceived strengths and Weaknesses in supporting ELT?*

The study highlighted the distinct strengths and weaknesses of ChatGPT, Diffit, and Brisk. While these tools could enhance teaching efficiency and student engagement, their limitations underscored the need for teacher expertise and reflective practices. Generative AI tools addressed key pedagogical challenges in innovative ways. ChatGPT's versatility supported differentiated instruction, enabling teachers to provide tailored materials for students at varying proficiency levels. Diffit's text simplification fostered inclusivity by making complex content accessible to lower-level learners. Brisk's multimedia presentations engaged students through visuals, supporting multimodal learning approaches. Despite their strengths, the tools often failed to deliver outputs that fully aligned with ELT objectives. ChatGPT's occasional inaccuracies and complex outputs required teacher intervention, while Diffit's oversimplification diminished the richness of particular texts. Brisk's shallow templates limited its use for advanced topics. These findings reflect the observations of Ghimire et al. (2024), who noted that while generative AI tools reduce teacher workload, they cannot replace the depth of human expertise in pedagogy.

The tools' limitations reinforce the need for ongoing professional development in AI literacy. Teachers must develop skills in prompt engineering and critical evaluation of AI outputs, as Nyaaba and Zhai (2024) suggested, to maximize these tools' effectiveness.

RQ3: *How do reflective practices help teachers adapt and overcome challenges using GAI?*

Reflective practices played a pivotal role in enabling teachers to effectively address the challenges associated with the integration of generative AI tools in language teaching. Teachers engaged in two key forms of reflection: *reflection-in-action*, which involved making immediate pedagogical decisions during the use of AI tools in the classroom, and *reflection-on-action*, which entailed evaluating and adjusting their instructional strategies after the teaching event. These reflective cycles were particularly evident during Focus Group Discussions (FGDs), where educators collaboratively analyzed their experiences. The group setting fostered shared learning and collective problem-solving, creating a supportive environment for professional growth.

This iterative reflective process resonates strongly with Schön's (1983) theory of Reflective Practice, which underscores the importance of critical, experience-based learning in professional contexts. Teachers exhibited a strong inclination toward *reflection-on-action*, as they critically examined the efficacy of AI-generated outputs, identified gaps, and revised their instructional approaches accordingly. One teacher, for instance, recounted refining her ChatGPT prompts after discovering that overly general queries produced off-topic results. This adjustment demonstrated the type of adaptive learning Schön envisioned—teachers solving pedagogical problems through reflective inquiry. Furthermore, the collaborative dimension of reflection was a notable outcome of the FGDs. Teachers shared insights, exchanged strategies,

and co-developed solutions for enhancing AI integration. This communal reflection aligns with the findings of Mishra et al. (2024), who emphasize the importance of teacher networks in fostering sustained professional development. Through dialogue and peer feedback, participants not only improved their individual practices but also cultivated a shared understanding of best practices for using generative AI tools effectively in the classroom.

Practical Implications for ELT Practitioners

The findings of this study carry significant implications for both theoretical development and practical application in the field of English Language Teaching (ELT). One key implication is the need to expand the Technological Pedagogical Content Knowledge (TPACK) framework to accommodate the integration of generative AI tools. This expansion should include AI-specific competencies such as prompt engineering and ethical considerations. As noted by Brandão et al. (2024), AI literacy ought to become a central element of teacher education programs to equip educators with the skills necessary to navigate the complexities of AI integration effectively.

Another important implication involves embedding reflective practices within professional development initiatives. The study highlighted the value of reflective thinking for addressing limitations in AI-generated materials. To foster ongoing growth and improvement, professional development should promote both reflection-in-action and reflection-on-action. Additionally, collaborative spaces like teacher-learning communities can enhance these efforts by enabling shared learning experiences and fostering innovation among educators.

Lastly, the study emphasizes the potential of generative AI tools to support differentiated instruction, allowing teachers to respond more effectively to diverse learner needs. However, the effective use of these tools depends on educators' ability to critically adapt AI outputs to ensure alignment with pedagogical and content objectives. This insight aligns with Zawacki-Richter et al. (2019), who argued that AI technologies should serve as enhancements to, rather than replacements for, the inherently human dimensions of teaching.

Generative AI tools like ChatGPT, Diffit, and Brisk promise to transform ELT by streamlining lesson preparation, enhancing accessibility, and engaging students. However, their limitations highlight the irreplaceable role of teachers in adapting these tools to meet pedagogical goals. Integrating AI into Education requires a dual focus on building technological competencies and fostering reflective practices, ensuring teachers remain central to the learning process. By situating these findings within the TPACK framework and Reflective Practice theory, this study contributes to a deeper understanding of how AI tools can support, rather than supplant, practical teaching.

CONCLUSION AND IMPLICATIONS
Conclusion

This study explored integrating generative AI tools—ChatGPT, Diffit, and Brisk—into English Language Teaching (ELT) by analyzing teachers' reflective journals, focus group discussions (FGDs), and document analysis. The findings revealed the tools' distinct use patterns, strengths, and limitations. ChatGPT was valued for its flexibility in generating diverse materials, Diffit excelled in simplifying complex texts to enhance accessibility, and Brisk engaged students with visually appealing multimedia presentations. However, all tools required significant teacher intervention to align outputs with instructional objectives.

The study highlights the critical role of reflective practices in enabling teachers to adapt and optimize the use of these tools. Teachers overcame challenges such as inaccuracies in AI-generated content, oversimplification, and limited depth through reflection-in-action, reflection-on-action, and collaborative reflections. These findings affirm the importance of teacher agency, technological literacy, and adaptive strategies in effectively integrating generative AI into ELT.

Grounded in the TPACK framework and Reflective Practice theory, this study contributes to the growing body of literature on AI in Education, offering practical insights into how generative AI tools can enhance, rather than replace, teachers' pedagogical expertise. While this study provides valuable insights, it also highlights areas for future research. Longitudinal studies could explore how teachers' AI literacy evolves and how sustained exposure to AI tools impacts pedagogical effectiveness and student learning outcomes. Additionally, while this study focused on teacher experiences, future research should examine student perspectives on AI-assisted learning, particularly concerning engagement, motivation, and language acquisition. Further investigation is needed into AI's ethical and equity considerations in ELT, particularly in addressing bias, misinformation, and data privacy concerns.

Implications of the Study
Theoretical Implications

The TPACK framework must evolve to include competencies specific to AI tools, such as prompt engineering, critical evaluation of AI outputs, and understanding the ethical considerations of AI in Education. This aligns with emerging research on teacher preparation for AI-supported classrooms (Mishra et al., 2024). The iterative process of reflection demonstrated by teachers highlights the value of Schön's Reflective Practice in technology integration. Future studies should explore how reflective practices can be formally embedded in teacher training programs to support AI adoption.

Practical Implications

Teachers need targeted professional development to maximize the potential of generative AI tools. Training programs should focus on 1) Crafting effective prompts for tools like ChatGPT, 2) Evaluating and editing AI-generated outputs to ensure alignment with pedagogical goals, and 3) Adapting tools like Diffit and Brisk to meet the needs of diverse learners. These skills will empower teachers to address the challenges associated with AI integration and optimize its benefits. Generative AI tools like Diffit demonstrate immense potential for differentiated instruction by simplifying content for lower-proficiency learners. However, teachers must combine these tools with supplementary materials to address critical details and higher-order thinking skills. Tools like Brisk can support multimodal learning by creating visually engaging content. Teachers should leverage these tools for introductory or discussion-based lessons, supplementing slides with additional depth where necessary.

Policy-Level Implications

Pre-service teacher training programs should incorporate AI literacy as a core component, ensuring that future educators can navigate and utilize generative AI tools effectively. This aligns with global calls for updating teacher education curricula to reflect advancements in AI technologies. Developers of generative AI tools must address ethical concerns, such as bias and cultural relevance, to ensure that outputs align with diverse classroom contexts. Policymakers should advocate for AI tools that are inclusive, unbiased, and adaptable to varying educational settings. Schools and academic institutions should create environments encouraging reflective practices and collaborative learning among teachers. Regular peer-review sessions, workshops, and communities of practice can foster shared knowledge and innovation in AI integration. This study highlights the transformative potential of generative AI tools in ELT while emphasizing that their success hinges on teacher expertise, reflective practices, and targeted professional development. By equipping teachers with the skills to navigate AI tools critically and creatively, education systems can harness the power of AI to enhance learning experiences and bridge gaps in accessibility, engagement, and differentiation.

REFERENCES

Albion, P., Jamieson-Proctor, R., & Finger, G. (2010). Auditing the TPACK competence and confidence of Australian teachers: The Teaching With ICT Audit Survey (TWictAS). *Society for Information Technology & Teacher Education International Conference (SITE)*, 1–8. http://eprints.usq.edu.au/7276/1/Albion_Jamieson-Proctor_Finger_SITE_2010_AV.pdf

Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal, 9*(2), 27–40. https://doi.org/10.3316/qrj0902027

Brandão, A., Pedro, L., & Zagalo, N. (2024). Teacher professional development for a future with generative artificial intelligence – An integrative literature review. *Digital Education Review, 45*, 151–157. https://doi.org/10.1344/der.2024.45.151-157

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Celik, I. (2023). Towards Intelligent-TPACK: An empirical study on teachers' professional knowledge to ethically integrate artificial intelligence (AI)-based tools into education. *Computers in Human Behavior, 138*, 107468.

Cogo, A., Patsko, L., & Szoke, J. (2024). Generative artificial intelligence and ELT. *ELT Journal, 78*(4), 373–377. https://doi.org/10.1093/elt/ccae051

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Sage Publications.

Dewi, F. (2019). *Technology-supported English language teaching: A professional development for EFL teachers' technology integration* [Master's thesis, Universitas Pendidikan Indonesia].

Dewi, F., & Hamied, F. A. (2024). Integration of advanced digital tools in language teaching and learning. In *Education in the Digital Era* (pp. 295–324). UPI Press.

Ghimire, A., Prather, J., & Edwards, J. (2024). Generative AI in education: A study of educators' awareness, sentiments, and influencing factors (Version 1). *arXiv*. https://doi.org/10.48550/ARXIV.2403.15586

Krushinskaia, K., Elen, J., & Raes, A. (2023). Systematic literature review on teacher's roles in AI-supported education. *Proceedings of the International Conference on Computer-Supported Collaborative Learning*, 370–371. https://doi.org/10.22318/cscl2023.109177

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record, 108*(6), 1017–1054.

Mishra, P., Warr, M., & Islam, R. (2023). TPACK in the age of ChatGPT and generative AI. *Journal of Digital Learning in Teacher Education, 39*(4), 235–251. https://doi.org/10.1080/21532974.2023.2247480

Morgan, D. L. (1997). *Focus groups as qualitative research* (2nd ed.). Sage Publications.

Novawan, A., Ikeda, O., & Walker, S. A. (2024). The new face of technology-enhanced language learning (TELL) with artificial intelligence (AI). [Publication details pending].

Nyaaba, M., & Zhai, X. (2024). Generative AI professional development needs for teacher educators. *Journal of AI, 8*(1), 1–13. https://doi.org/10.61969/jai.1385915

Patton, M. Q. (2002). *Qualitative research & evaluation method* (3rd ed.). Sage Publications. https://doi.org/10.1016/j.ijproman.2012.11.012

Richards, J. C., & Farrell, T. S. C. (2011). *Practice teaching: A reflective approach* (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781139151535

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. Basic Books.

Stake, R. E. (2010). *Qualitative research: Studying how things work*. The Guilford Press.

Susanto, D. A., Priyolistiyanto, A., Pinandhita, F., K.A, A. B. P., & Bimo, D. S. (2024). Utilizing ChatGPT on designing English language teaching (ELT) materials in Indonesia: Opportunities and challenges. *Celt: A Journal of Culture, English Language Teaching & Literature, 24*(1), 157–171. https://doi.org/10.24167/celt.v24i1.11633

Tai, S.-J. D. (2015). From TPACK-in-action workshops to classrooms: CALL competency developed and integrated. *Language Learning & Technology, 19*(1), 139–164.

Tunjera, N., & Chigona, A. (2023). Investigating effective ways to use artificial intelligence in teacher education. *European Conference on E-Learning, 22*(1), 331–340. https://doi.org/10.34190/ecel.22.1.1625

Yin, R. K. (2009). *Case study research: Design and methods* (2nd ed.). SAGE Publications.

Zawacki-Richter, 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*(1), 39. https://doi.org/10.1186/s41239-019-0171-0