RESEARCH ARTICLE VOLUME 19, ISSUE 2, 44-58

The New English Teacher
ISSN 2985-0959 (Online)

Towards Technology-Enhanced English Learning: Gender Analysis of AI Large Language Models (LLMs) Usage

Budi Waluyo
School of Languages and General Education
Walailak University
222 Thai-Buri sub-district, Thasala District,
Nakhon Si Thammarat, Thailand, 80160
Tel: +66093-036-9427
Email: budi.business.waluyo@gmail.com

Adi Isma
Faculty of Teachers Training and Education
Universitas Sulawesi Barat
Jalan Prof. Dr. Baharuddin Lopa, Baurung,
Banggae Timur, Majene Regency, West
Sulawesi, Indonesia, 91412
Tel: +62-8524-215-7825
Email: adi.isma@unsulbar.ac.id

Abstract: Despite the growing integration of Artificial Intelligence (AI) Large Language Models (LLMs) in education, research exploring gender-specific perceptions and usage remains scarce. This study employed the Unified Theory of Acceptance and Use of Technology (UTAUT) and a mixed-methods sequential explanatory design to investigate the perceptions and usage of ChatGPT among 41 female and 41 male university students for English learning. Data were collected using a technology acceptance scale and structured written interviews and analyzed through descriptive and inferential statistics (t-tests, bivariate correlations, and one-way ANOVA) for quantitative data, and thematic analysis for qualitative insights. Findings indicate no significant gender differences in the overall acceptance of ChatGPT. However, gender-specific trends were observed: males predominantly valued ChatGPT for specific academic tasks, whereas females appreciated its general enhancement of their learning experience. English proficiency levels did not notably affect perceptions of ChatGPT's utility, suggesting consistent recognition of its benefits across varying proficiency levels, thereby refuting previous assumptions about proficiency impacting technology adoption. This study highlights the roles that gender and proficiency level play in shaping the educational use of AI language models.

Keywords: gender, artificial intelligence, ChatGPT, language proficiency

Received: July 7, 2024 Revised: January 2, 2025 Accepted: March 20, 2025

Introduction

The integration of Artificial Intelligence (AI) Large Language Models (LLMs) into education has redefined foreign language learning by offering personalized instruction, real-time feedback, and extensive linguistic resources (Bahroun et al., 2023; Morosanu et al., 2023; Zaim et al., 2024). Exemplified by ChatGPT, such technologies have been praised for enhancing student engagement and linguistic proficiency (Gayed et al., 2022; Ranalli, 2021); however, a significant gap remains in understanding the intersection of gender and language proficiency in AI adoption within higher education (Jeon et al., 2023). While research has largely concentrated on primary and secondary education, uncovering gender-specific differences in technology use that may exacerbate

disparities in learning outcomes (Xia et al., 2023; Zhang et al., 2023), and the implications for university-level learning are underexplored. The lack of focus on tertiary education contexts overlooks the need for inclusive pedagogical strategies that address diverse learner needs and ensure equitable access to AI-driven innovations. Investigating gendered perceptions and usage patterns of AI LLMs in higher education can illuminate pathways for developing adaptive teaching approaches and fostering inclusivity in educational technologies.

The advent of AI LLMs, epitomized by ChatGPT, marks a transformative leap in natural language processing, offering unprecedented human-like text generation capabilities. Released by OpenAI in November 2022, ChatGPT attracted over 100 million subscribers within two months, reflecting its widespread appeal and utility (Hu, 2023). By facilitating intuitive human-computer interactions and advancing machines' ability to comprehend and replicate human language with remarkable precision, the model has redefined the benchmarks of natural language processing (Nah et al., 2023; Kuhail et al., 2023). Research accentuates its multifaceted applications in English language learning, particularly in higher education, where it enhances various skills. AI LLMs improve writing by providing instant feedback on grammar, vocabulary, and structure, enabling students to produce clearer, error-free texts (Malik et al., 2023; Song & Song, 2023). Speaking proficiency benefits from voice recognition technologies that deliver personalized feedback on pronunciation and fluency (Rasul et al., 2023), while contextualized exercises and active listening strategies generated by these models support reading and listening comprehension, enhancing retention and understanding of texts and spoken language (Baskara, 2023; Li et al., 2023; Zhang & Huang, 2024).

Nonetheless, the integration of AI technologies in education presents sophisticated challenges, particularly when examined through the lens of gender dynamics. Although socio-cultural, educational, and economic shifts have reduced historical gender disparities in technology adoption, significant gaps persist in professional domains such as software development and AI due to entrenched biases, limited mentorship, socioeconomic barriers, and societal expectations that hinder female participation in tech-related fields (Eckhardt et al., 2021; Trauth & Connolly, 2021). Research has further highlighted the gender bias inherent in tools like ChatGPT, which may exacerbate inequalities and disadvantage individuals across gender identities (Gross, 2023). Language proficiency adds another layer of complexity, as individuals with advanced proficiency often derive greater benefits from interactions with AI LLMs, including more effective communication and learning outcomes, while those with lower proficiency face difficulties in query formulation and response comprehension, limiting their educational gains (Kuhail et al., 2023; Shaikh et al., 2023). The intersection of gender and language proficiency further compounds these challenges, as societal norms and gendered socialization influence language acquisition and consequently affect the use and benefits of AI tools in educational contexts (Li & Xu, 2020; Malik et al., 2023; Strzelecki, 2023).

Therefore, this study intends to address the critical gap in understanding gender-specific perceptions and usage patterns of ChatGPT among university students engaged in English language learning, highlighting its significance in fostering inclusive and equitable educational technologies. Unlike existing research predominantly focused on primary and secondary education, this study focuses on the unique dynamics of tertiary education, where gendered interactions with AI tools present distinct challenges and opportunities. By examining how gender influences the acceptance and utilization of ChatGPT, the research stresses the importance of addressing disparities to ensure equitable access and meaningful engagement for diverse learner groups. Grounded in the Unified Theory of Acceptance and Use of Technology (UTAUT), this study employs a theoretical framework to explore key predictors of technology acceptance, such as performance expectancy, effort expectancy, social influence, and facilitating conditions, within gendered contexts. The rationale lies in the urgent need to align AI adoption with principles of inclusivity, as understanding these

dynamics can guide the development of more responsive and adaptable educational technologies. Specifically, this study seeks to investigate:

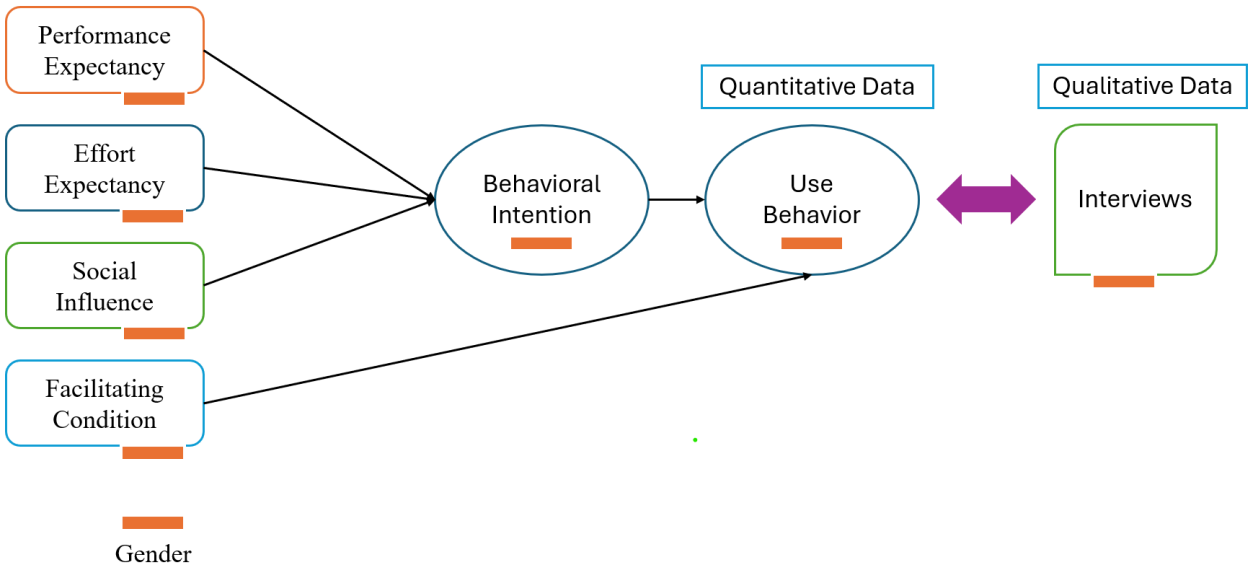
1. What are the differences in perceptions and usage patterns of ChatGPT between male and female university students in learning English?
2. How do gender-specific behaviors correlate with the adoption and effectiveness of ChatGPT in English language learning?
3. To what extent do English proficiency levels moderate the relationship between gender and ChatGPT usage in an academic context?

Methodology

Mixed-Methods Sequential Explanatory Design

The mixed-methods sequential explanatory design efficiently combines quantitative and qualitative research in two phases to enrich data analysis. Initially, it involves quantitative data collection and analysis to identify students' voices and interactions with *ChatGPT*. This is followed by qualitative research, using written interviews in depth to examine aspects such as gender differences in AI use and perception. Figure 1 below illustrates the research design, adapted from Yilmaz et al. (2023).

Figure 1
Illustration of the Research Design



Research Context and Participant

This study utilized a convenience sampling method and was conducted at several universities across Sulawesi Island, Indonesia, involving 82 undergraduate students (50% male and 50% female) who all major in various fields of education, including English, Childhood, Mathematics, and Science Education. The participant group comprised an equal number of male and female students, each subgroup containing 41 individuals. The female students had an average age of 20.15 years (SD = 1.98) with English proficiency levels distributed as 31.7% poor, 46.3% moderate, and 22% good. Conversely, the male students were slightly older, with an average age of 20.78 years (SD = 3.81), and their English proficiency was reported as 41.5% poor, 48.8% moderate, and 9.8% good. All participants confirmed familiarity with AI language model applications like ChatGPT, indicating a baseline of engagement with digital language tools.

Data Collection and Research Instrument

In this study, quantitative data was gathered using an adapted acceptance scale from Yilmaz et al. (2023), which used a 1 to 5 rating system to gauge levels of agreement. The scale featured 18 items, organized into four sub-scales: Performance Expectancy, Effort Expectancy, Facilitating Conditions, and Social Influence, aiming to examine various dimensions of technology acceptance. Behavioral Intention was conceptualized as a construct influenced by Performance Expectancy, Effort Expectancy, and Social Influence, while Use Behavior was analyzed through the collective lens of all four sub-scales, as illustrated in Figure 1.

To qualitatively validate the instrument items, a distinguished English Language Teaching (ELT) expert, with a specialization at the nexus of ELT and technology, conducted a face validity assessment. Further scrutiny into the scale's reliability through internal consistency checks yielded Cronbach's alpha values above .70 across all sub-scales, affirming their reliability. Notably, the scale's overall internal consistency was exceptionally high (α = .96), leading researchers to maintain its comprehensive structure for analysis, ensuring a robust examination of technology acceptance behaviors among participants, as presented in Table 1.

Table 1
Construct and Internal Reliability Results

Constructs and Sample Statements	***α***
1. Performance Expectancy	.95
2. Effort Expectancy	.91
3. Facilitating Condition	.90
4. Social Influence	.93
4. Behavioral Intention	.95
5. Use Behavior	.96

To quantitatively validate the questionnaire used in this study, the researchers employed the KMO and Bartlett's tests, which yielded significant results ($\chi^2(190) = 1466.92$, $p < .001$) and demonstrated a sampling adequacy of .904, well above the commonly accepted threshold of .50. These results affirm that the survey constructs are robust and effectively capture the intended information. Additionally, the data conformed to a normal distribution, as evidenced by skewness and kurtosis values within the ±2 range. This normalcy in data distribution validates the use of parametric tests in subsequent stages of data analysis, ensuring the statistical integrity and reliability of the findings derived from this survey data.

For the qualitative component of the study, structured written interviews with all participants were used to explore how students utilize ChatGPT for English learning in higher education contexts. A set of interview questions was created in Google Forms and distributed online to all participants via links. Such a method was specifically chosen to mitigate any potential misunderstandings caused by language barriers between the researchers and participants. It was also deemed effective in accommodating the students' hectic schedules and the logistical challenges posed by long distances, in line with the protocols suggested by Jacob and Furgerson (2012). To maintain anonymity and organize data efficiently, students were assigned codes based on their gender: 'F' for females (e.g., F1, F2, etc.) and 'M' for males (e.g., M1, M2, etc.), facilitating a systematic analysis of the responses.

Data Analysis

The analysis of the survey data entailed two primary steps: categorizing responses for each subscale into three levels—low engagement (0 to 1.7), moderate engagement (1.8 to 3.4), and high engagement (3.5 to 5.0) for descriptive statistics; and employing inferential statistical methods to deepen the understanding of the data. These methods included independent t-tests to compare group means, bivariate correlations to explore the relationships among variables, and one-way ANOVA to evaluate differences across multiple groups. Additionally, qualitative data underwent thematic analysis following the procedures outlined by Clarke and Braun (2017), which facilitated the identification, analysis, and reporting of emergent themes, thereby offering profound insights into student interactions with ChatGPT in their English learning within higher education contexts. The coding of qualitative data was conducted manually by a single researcher, adhering to rigorous and systematic methods recommended by Clarke and Braun (2017), rendering intercoder agreement non-applicable in this context.

Results

Quantitative Findings

Descriptive Statistics Analysis

In our analysis of undergraduate students' perceptions of using *ChatGPT* for English learning, we investigated several variables, including Performance Expectancy, Effort Expectancy, Facilitating Conditions, Behavioral Intention, and Use Behavior. These factors were quantitatively assessed through mean scores and standard deviations to determine the levels of acceptance—high, moderate, or low—among male and female students. The results reveal that both genders perceive *ChatGPT* as a beneficial tool for enhancing English proficiency. Performance Expectancy was highly rated by both genders, with males (M = 3.6, SD = 0.9) showing slightly more variation in their responses than females (M = 3.6, SD = 0.7). Females rated the ease of using *ChatGPT* (Effort Expectancy) higher (M = 3.6, SD = 0.6) than males (M = 3.6, SD = 0.9), and both agreed on its facilitating conditions, indicating good support and resources for effective use (Males: M = 3.5, SD = 0.9; Females: M = 3.5, SD = 0.8).

Furthermore, the study examined students' Behavioral Intentions and Use Behaviors to understand their plans and actual usage of ChatGPT. High enthusiasm was noted across genders, with females showing stronger intentions (M = 3.6, SD = 0.6) and slightly more consistent use (M = 3.5, SD = 0.6) compared to males (Intentions: M = 3.6, SD = 0.8; Use: M = 3.5, SD = 0.7). These findings highlight not only the general acceptance of *ChatGPT* as a learning aid but also the nuanced differences in usage patterns between male and female students, which could suggest varying degrees of reliance on technological aids in language acquisition. The detailed statistics are consolidated in Table 2, providing a clear comparison of how male and female students utilize *ChatGPT* in their English language studies.

Table 2
Descriptive Statistics on the Students' Acceptance and Use Levels of ChatGPT

	Males			Females		
	M	**SD**	**Level**	**M**	**SD**	**Level**
Performance Expectancy	3.6	0.8	High	3.6	.7	High
Effort Expectancy	3.6	0.9	High	3.6	.6	High
Facilitating Conditions	3.5	0.9	High	3.5	.8	High
Social Influence	3.3	0.8	Moderate	3.3	.7	Moderate
Behavioral Intentions	3.6	0.8	High	3.6	.6	High
Use Behaviors	3.5	0.7	High	3.5	.6	High

Independent T-Tests

The independent t-tests comparing male and female undergraduate students' acceptance levels toward using ChatGPT in English language learning revealed no significant gender differences across several evaluative categories, including Performance Expectancy, Effort Expectancy, and Facilitating Conditions. Specifically, Performance Expectancy showed a t-value of .12 and a p-value of .90, indicating a similar recognition of *ChatGPT*'s potential to enhance learning with a negligible mean difference between genders (mean difference = .02, 95% CI = -.32 to .36). The analysis for Effort Expectancy and Facilitating Conditions further supported this gender-neutral perspective, with t-values of -.32 and -.04, and p-values of .75 and.97, respectively. These *p*-values, well above the significance level of .05, underline a consensus on the ease of use and the adequacy of resources supporting *ChatGPT*'s integration into learning processes, regardless of gender.

Further statistical testing on Behavioral Intention and Use Behavior also accentuated similar attitudes toward future and actual use of *ChatGPT* between male and female students. The Behavioral Intention category revealed a t-value of -.09 and a *p*-value of .93, with a mean difference of -.01, suggesting a shared likelihood and uniform expectation of employing *ChatGPT* as a learning tool. Meanwhile, Use Behavior exhibited a t-value of -.08 and a p-value of .94 (mean difference = -.01, 95% CI = -.30 to .28), highlighting no substantial difference in actual usage levels. These findings, as seen in Table 3, collectively emphasize the equitable acceptance and application of *ChatGPT* among students, affirming its utility and inclusive appeal in educational settings without gender bias.

Table 3
Independent T-Test Results

	F	t	p	Mean Difference	95% Confidence Interval	
					Lower	Upper
Performance Expectancy	1.09	.12	.90	.02	-.32	.36
Effort Expectancy	3.50	-.32	.75	-.05	-.39	.28
Facilitating Conditions	.23	-.04	.97	-.01	-.37	.36
Social Influence	1.74	-.03	.98	-.01	-.34	.33
Behavioral Intention	.00	-.09	.93	-.01	-.32	.29
Use Behavior	.27	-.08	.94	-.01	-.30	.28

Correlations

The correlation analysis of undergraduate students' acceptance and use of *ChatGPT* for English language learning reveals distinct patterns in how various factors interrelate, particularly between different expectancy measures and their overall behavioral intentions and use behaviors, as indicated in Table 4. For males, Performance Expectancy (PE) shows strong positive correlations with all other measures: .597 with Effort Expectancy (EE), .558 with Facilitating Conditions (FC), .700 with Social Influence (SC), and notably high correlations with Behavioral Intention (BI) and Use Behavior (UB) at .828 and .847 respectively. These correlations suggest that males who anticipate beneficial outcomes from ChatGPT use (high PE) are also those who find it easier to use (high EE). They believe they have the requisite support to use it (high FC), feel social support for its use (high SC), and consequently show higher intentions to use and actual use of *ChatGPT* (high BI and UB).

For females, the pattern is similarly strong but features even higher correlation coefficients, indicating slightly tighter interdependencies among the measures: PE correlates with EE at .835, with FC at .801, and with SC at a lower but significant .410, while correlating extremely strongly with BI and UB at .942 and .890 respectively. The Effort Expectancy for females also shows strong

positive correlations with Facilitating Conditions (.769), Social Influence (.490), and exceptionally strong ones with Behavioral Intention (.922) and Use Behavior (.898). These robust correlations, as shown in Table 3, reinforce the perceived ease of use and the supportive conditions as crucial for females' high engagement levels with *ChatGPT*. Across both genders, the data highlights a critical insight: higher expectations and better facilitation conditions not only enhance students' intentions to use *ChatGPT* but are closely linked to their actual usage behaviors, illustrating a dynamic interplay between cognitive assessments and practical engagement with technology in educational settings.

Table 4

Correlation Analysis Results

	Males					Females				
	EE	**FC**	**SC**	**BI**	**BU**	**EE**	**FC**	**SC**	**BI**	**UB**
Performance Expectancy (PE)	.597**	.558**	.700**	.828**	.847**	.835**	.801**	.410**	.942**	.890**
Effort Expectancy (EE)		.682**	.550**	.885**	.849**		.769**	.490**	.922**	.898**
Facilitating Conditions (FC)			.528**	.870**	.832**			.517**	.930**	.912**
Social Influence (SC)				.685**	.822**				.509**	.711**
Behavioral Intention (BI)					.978**					.967**

Note. ** Correlation is significant at the 0.01 level (2-tailed).

One-Way ANOVA

The one-way ANOVA results, as seen in Table 5, reveal that there are no significant differences within male or female groups for the various constructs related to the acceptance of *ChatGPT* for English language learning. For both genders, all constructs—Performance Expectancy, Effort Expectancy, Facilitating Conditions, Social Influence, Behavioral Intention, and Use Behavior—yield high p-values (well above the significance threshold of 0.05), indicating the differences between groups are statistically insignificant. For instance, Performance Expectancy among males (F = .52, p = .60) and females (F = .35, p = .71) and Effort Expectancy among males (F = .12, p = .89) and females (F = .50, p = .61) show that the perceptions within each gender are uniform and do not vary significantly based on the groups considered in this study. The most noteworthy aspect of these results is the uniformity of the findings across all constructs and both genders, pointing to a cohesive perception of *ChatGPT*'s utility in language learning. Even for Social Influence, where the between-groups variance is relatively larger for females (F = 1.74, p = .19), it does not reach statistical significance, further confirming the homogeneity within the groups. Overall, this lack of significant difference underlines the widespread and consistent acceptance of *ChatGPT* across the cohorts sampled, suggesting that any variations in acceptance levels are not due to random group differences but potentially to other factors not captured in this analysis.

Table 5
One-Way ANOVA Results

		Males				Females			
		Sum of Squares	**Mean Square**	**F**	**p**	**Sum of Squares**	**Mean Square**	**F**	**p**
Performance Expectancy	Between Groups	.76	.38	.52	.60	.33	.16	.35	.71
	Within Groups	28.09	.74			17.61	.46		
Effort Expectancy	Between Groups	.21	.10	.12	.89	.37	.18	.50	.61
	Within Groups	32.10	.85			13.86	.37		
Facilitating Conditions	Between Groups	.62	.31	.37	.69	.03	.02	.03	.97
	Within Groups	31.63	.83			22.66	.60		
Social Influence	Between Groups	1.35	.67	.97	.39	1.55	.78	1.74	.19
	Within Groups	26.34	.69			16.95	.45		
Behavioral Intention	Between Groups	.24	.12	.20	.82	.19	.09	.23	.80
	Within Groups	22.86	.60			15.54	.41		
Use Behavior	Between Groups	.33	.17	.30	.74	.03	.02	.04	.96
	Within Groups	20.89	.55			13.22	.35		

Furthermore, the Tukey HSD post-hoc test results, as presented in Table 6, for comparing proficiency levels among male and female students yield insights that are quite revealing. The findings consistently report high p-values across all categories, suggesting no statistically significant differences in the mean scores for any of the proficiency comparisons within both male and female groups. For males, the comparison between the poor and moderate proficiency levels in Performance Expectancy (PE) resulted in a *p*-value of .95, while the comparison between poor and good proficiency levels showed a *p*-value of .77, indicating no significant differences in how male students of different proficiencies perceive the effectiveness of *ChatGPT* in enhancing their learning. This pattern holds true across other constructs such as Effort Expectancy (EE), Facilitating Conditions (FC), Social Influence (SC), Behavioral Intention (BI), and Use Behavior (UB) with *p*-values such as .99 for EE, .92 for FC, and .97 for UB, which reinforces the non-significant differences.

For females, the trend is analogous, as evidenced by the comparisons in PE, which yielded p-values of .87 and .88 for moderate versus poor and good versus poor, respectively. In other areas, such as Social Influence, even when the mean difference between poor and moderate proficiency seems substantial (mean difference = .40), the significance level (p = .18) does not suggest a statistically relevant distinction. The consistent lack of significant differences within the groups for both males and females implies that proficiency levels do not affect students' perceptions and interactions with *ChatGPT* in a statistically meaningful way. This uniformity suggests a broad and homogeneous acknowledgment of *ChatGPT*'s potential benefits and utility across various proficiency levels, indicating that student acceptance and expected use of *ChatGPT* for English language learning are likely not influenced by their current proficiency.

Table 6
Results of the Tukey HSD

			Males			Females		
Dependent Variable	**(I) proficiency**	**(J) proficiency**	**Mean Difference (I-J)**	**Std. Error**	**p**	**Mean Difference (I-J)**	**Std. Error**	**p**
PE	Poor	Moderate	.09	.31	.95	-.11	.23	.87
		Good	-.26	.37	.77	.18	.38	.88
	Moderate	Poor	-.09	.31	.95	.11	.23	.87
		Good	-.35	.35	.57	.29	.37	.71
	Good	Poor	.26	.37	.77	-.18	.38	.88
		Moderate	.35	.35	.57	-.29	.37	.71
EE	Poor	Moderate	-.16	.33	.88	-.18	.20	.65
		Good	-.06	.40	.99	.05	.34	.99
	Moderate	Poor	.16	.33	.88	.18	.20	.65
		Good	.10	.37	.96	.23	.33	.77
	Good	Poor	.06	.40	.97	-.05	.34	.99
		Moderate	-.10	.37	.96	-.23	.33	.77
FC	Poor	Moderate	.13	.33	.92	-.01	.26	1
		Good	-.19	.40	.88	.09	.43	.97
	Moderate	Poor	-.13	.33	.92	.01	.25	1
		Good	-.32	.37	.67	.1	.43	.97
	Good	Poor	.19	.40	.88	-.09	.43	.97
		Moderate	.32	.37	.67	-0.1	.42	.97
SC	Poor	Moderate	-.33	.30	.53	.40	.22	.18
		Good	-.47	.36	.41	.06	.37	.99
	Moderate	Poor	.33	.30	.53	-.40	.22	.18
		Good	-.14	.34	.91	-.34	.37	.63
	Good	Poor	.47	.36	.41	-.06	.37	.99
		Moderate	.14	.34	.91	.34	.37	.63
BI	Poor	Moderate	.02	.28	.10	-.10	.21	.89
		Good	-.17	.34	.87	.11	.36	.95
	Moderate	Poor	-.02	.28	.10	.10	.21	.89
		Good	-.19	.31	.82	.21	.35	.83
	Good	Poor	.17	.34	.87	-.11	.36	.95
		Moderate	.19	.31	.82	-.21	.35	.83
UB	Poor	Moderate	-.07	.27	.97	.03	.20	.99
		Good	-.25	.32	.73	.10	.33	.95
	Moderate	Poor	.07	.27	.97	-.03	.20	.99
		Good	-.18	.30	.82	.07	.32	.97
	Good	Poor	.25	.32	.73	-.10	.33	.95
		Moderate	.18	.30	.82	-.07	.32	.97

Qualitative Findings

Table 7
Emerging Themes about the Students' Use of ChatGPT

Themes	Males	Keywords	Females	Keywords
1	Assistance with Assignments	assistance, assignments, references, plagiarism	Versatile Learning Tool	versatile, grammar checks, vocabulary, comprehension
2	Learning Aid for English	learning aid, English, vocabulary, grammar	Personalized Learning	personalized learning, tailored materials, individual needs
3	Convenience and Accessibility	convenience, accessibility, user-friendly, quick access	Caution against Over-Reliance	over-reliance, supplementation, optimal results
4	Learning Enhancement	learning enhancement, understanding, examples, improvement	Learning Engagement	engagement, interactive practice, exploration

The thematic analysis detailed in Table 7, illustrating how male and female students utilize ChatGPT in their university-level English learning, offers significant insights that are effectively contextualized within the Unified Theory of Acceptance and Use of Technology (UTAUT). For male students, the technology primarily serves as a pragmatic aid in navigating academic assignments, illustrating the UTAUT's performance expectancy facet, which predicts technology's acceptance based on its perceived benefits. Nevertheless, some male students expressed concerns about potential over-reliance on the AI, fearing it could lead to academic dishonesty issues like plagiarism. This apprehension highlights the need for a balanced approach to integrating AI tools in education, suggesting that while they can enhance academic performance, they also require users to engage critically to avoid dependency, which could undermine genuine learning and academic integrity. Below are the sample responses:

> Using *ChatGPT* to assist with college assignments is quite helpful, although sometimes there is an excessive focus on word usage, which might lead to plagiarism due to the habit of copying and pasting directly from the application. (M1)

> When I encounter unknown topics while learning, I use *ChatGPT* because it greatly facilitates the learning process and enhances my understanding of the material. (M19)

> *ChatGPT* assists in English language learning at university by providing additional examples, making corrections, offering suggestions, and improving texts. (M26)

In contrast, female students appreciate ChatGPT for its versatility and adaptability, aligning with UTAUT's effort expectancy dimension, which focuses on the ease of use and user-friendliness of technology. They value the platform's capacity to personalize learning experiences and provide support in areas like grammar and vocabulary, emphasizing the importance of technology that adapts to diverse learning needs. Moreover, female students' cautious approach to relying on ChatGPT and their advocacy for a blend of AI and traditional learning resources reflect the social influence component of UTAUT. This perspective points out the significance of external influences, such as societal norms and educational frameworks, in shaping technology adoption and utilization. The following are the sample responses:

> I use it for grammar checks and to find meanings of words I don't understand. (F8)
>
> *ChatGPT* is handy for finding brief explanations and answering questions, making tasks easier. However, it might discourage creative thinking. (F34)
>
> *ChatGPT* is very helpful for learning English as it allows for the exploration of vocabulary in context and offers interactive practice. (F41)

Both male and female groups acknowledge ChatGPT's user-friendly interface and its role in simplifying complex educational tasks, pointing to the facilitating conditions of UTAUT. These conditions suggest that the presence of supportive and enabling environments is crucial for the successful adoption of technology. ChatGPT's instant feedback and ability to assist with a variety of learning tasks facilitate a more engaging and effective educational experience. The emphasis on the platform's role in enhancing understanding and offering tailored learning aids indicates that when integrated thoughtfully, technology can significantly improve educational outcomes by aligning with students' individual learning preferences.

Overall, the thematic analysis revealed how gender dynamics influenced the acceptance and utilization of AI technologies like ChatGPT, accentuating distinct usage patterns and perceptions among male and female students. Male students often emphasized practical benefits, such as improved efficiency in completing assignments, while expressing concerns about over-reliance and plagiarism, suggesting a need for strategies that promoted critical engagement with AI tools. Female students, on the other hand, valued ChatGPT for its adaptability and personalized support in areas like grammar and vocabulary, advocating for its integration alongside traditional learning methods. These findings indicate the importance of gender-sensitive educational strategies that address specific concerns and maximize the tool's benefits.

Discussion and Implications

The study sought to investigate gender-specific perceptions and usage patterns of *ChatGPT* among university students learning English, addressing the underexplored intersection of gender, language proficiency, and AI adoption in higher education. Even though prior research has focused predominantly on primary and secondary education, exposing gender disparities in technology adoption and learning outcomes (Xia et al., 2023; Zhang et al., 2023), the implications for tertiary education remain inadequately examined. *ChatGPT*, with its capacity to provide personalized instruction, real-time feedback, and vast linguistic resources (Bahroun et al., 2023; Morosanu et al., 2023; Waluyo & Kusumastuti, 2024), offers a transformative opportunity for language learners. The findings of the current study demonstrate widespread acceptance of *ChatGPT* across genders, with both male and female students perceiving it as an effective tool for enhancing English proficiency. Nevertheless, the slightly higher ratings among females for effort expectancy and behavioral intention suggest a greater inclination to integrate *ChatGPT* into their learning routines. These results align with Li and Xu (2020), who highlighted the influence of gendered socialization on technology adoption, but they also challenge traditional narratives of significant gender disparities, indicating that technological inclusivity has improved in higher education.

The correlation analyses highlight the interplay between cognitive assessments and behavioral engagement with *ChatGPT*, revealing critical gender-based nuances. Strong correlations between performance expectancy, effort expectancy, and facilitating conditions with behavioral intention and use behavior indicate that perceived benefits and available resources significantly shape students' interactions with the tool. Nonetheless, females exhibited even stronger correlations, particularly between effort expectancy and behavioral intention, suggesting a higher dependency on user-friendly interfaces and supportive features. This finding resonates with Morosanu et al.

(2023), who emphasized the role of adaptive and accessible technologies in fostering engagement, and it stresses the need to address how gendered expectations and preferences shape technology adoption. For male students, the robust correlations between performance expectancy and behavioral outcomes reflect a results-oriented approach, where tangible benefits, such as improved academic performance, are primary motivators. This observation aligns with Kuhail et al. (2023), who argued that aligning technological advantages with users' expectations is pivotal for sustained engagement. The gender differences in cognitive and behavioral patterns suggest that *ChatGPT*'s design successfully caters to diverse user needs but highlights areas for further improvement in ensuring equitable engagement.

The thematic analysis provides deeper insights into how male and female students perceive and utilize *ChatGPT* in their learning processes. Male students highlighted its utility for assignments and simplifying complex tasks, often expressing concerns about over-reliance and potential academic integrity issues, such as plagiarism. These concerns mirror Gross's (2023) critique of the unintended consequences of AI integration, including reduced critical thinking and creativity. Female students, by contrast, emphasized *ChatGPT*'s adaptability and personalization, praising its ability to assist with grammar, vocabulary, and comprehension tailored to individual needs. This aligns with Malik et al. (2023), who stressed the importance of integrating AI tools into pedagogical frameworks to complement rather than replace traditional methods. The emphasis on engagement and exploration among female students further supports findings by Rasul et al. (2023), who noted that interactive features in AI tools can enhance motivation and foster active learning. The contrast in perspectives features the importance of balancing AI adoption with strategies that mitigate risks, such as over-reliance, while amplifying the benefits of personalization and interactivity.

The uniformity in perceptions across proficiency levels, as evidenced by the lack of significant differences in the ANOVA and post-hoc tests, highlights ChatGPT's potential to transcend linguistic barriers and provide consistent benefits to diverse learners. Another interpretation is that the differences in language proficiency between male and female student groups could act as a confounding variable, potentially impacting the results. Future research should consider additional qualitative exploration, such as interviews, to gain insights into how proficiency influences ChatGPT usage within each gender group. Moreover, qualitative responses reveal that lower-proficiency students may face challenges in fully leveraging the tool's potential, particularly in constructing effective queries and interpreting complex responses. This observation aligns with Strzelecki (2023), who emphasized the importance of scaffolding to support underprepared learners engaging with advanced AI tools. Furthermore, the ability of *ChatGPT* to provide contextualized exercises and adaptive feedback aligns with Baskara (2023) and Zhang and Huang (2024), who highlighted the effectiveness of AI tools in enhancing comprehension and retention. Addressing the challenges faced by less proficient learners through targeted support and user training can further optimize *ChatGPT*'s inclusivity and efficacy (Waluyo, 2024). By aligning its design and implementation with the principles of equitable access and adaptability, *ChatGPT* holds the potential to become a cornerstone in reshaping language education in higher education contexts.

Conclusion, Limitation, and Recommendation

The research delineates the extensive recognition and the significant advantages *ChatGPT* delivers to university students in enhancing English language proficiency, where participants from both genders recognized its efficacy in enhancing linguistic skills, offering personalized feedback, and increasing engagement. Although discernible gender disparities were minimal, female respondents exhibited a marginally higher dependence on the platform's user-friendly functionalities, highlighting the critical necessity for AI tools to be both adaptive and inclusive. The study, nonetheless, is constrained by its dependence on self-reported measures, which might not accurately reflect actual usage behaviors or comprehensively identify the specific challenges

encountered by learners with lower proficiency. Besides, the small sample size may have limited the ability to generalize the study's findings to a broader population and, to some extent, may not have revealed hidden differences, especially in the statistical analysis employed.

However, the study followed standard empirical procedures for both qualitative and quantitative data analysis, and the findings should be considered acceptable. We encourage other researchers to conduct similar studies to assess the reliability of these findings, although we have already compared our results with previous research in an earlier section. Furthermore, the absence of longitudinal data limits the ability to ascertain the enduring effects of *ChatGPT* on educational outcomes and its capacity to redress or perpetuate educational disparities. Subsequent investigations should consider the prolonged influences of AI on language development, evaluate interaction patterns across varied cultural and educational landscapes, and develop methods to mitigate dependency while promoting innovation and analytical thinking. It is recommended that educators and technology developers implement support mechanisms and offer precise training for learners with lower proficiency to guarantee equitable access, optimize educational benefits, and ensure the judicious use of *ChatGPT* as a transformative tool in education.

References

Bahroun, Z., Anane, C., Ahmed, V., & Zacca, A. (2023). Transforming education: A comprehensive review of generative artificial intelligence in educational settings through bibliometric and content analysis. *Sustainability*, *15*(17), 12983. https://doi.org/10.3390/su151712983

Baskara, R. (2023). Exploring the implications of ChatGPT for language learning in higher education. *Indonesian Journal of English Language Teaching and Applied Linguistics*, *7*(2), 343–358.

Clarke, V., & Braun, V. (2017). Thematic analysis. *The Journal of Positive Psychology*, *12*(3), 297–298.

Eckhardt, J., Kaletka, C., Pelka, B., Unterfrauner, E., Voigt, C., & Zirngiebl, M. (2021). Gender in the making: An empirical approach to understand gender relations in the maker movement. *International Journal of Human-Computer Studies*, *145*, 102548. https://doi.org/10.1016/j.ijhcs.2020.102548

Gayed, J. M., Carlon, M. K. J., Oriola, A. M., & Cross, J. S. (2022). Exploring an AI-based writing assistant's impact on English language learners. *Computers and Education: Artificial Intelligence*, *3*, 100055. https://doi.org/10.1016/j.caeai.2022.100055

Gross, N. (2023). What ChatGPT tells us about gender: A cautionary tale about performativity and gender biases in AI. *Social Sciences*, *12*(8), 435. https://doi.org/10.3390/socsci12080435

Hu, G. (2023). Challenges for enforcing editorial policies on AI-generated papers. *Accountability in Research*, 1–3. https://doi.org/10.1080/08989621.2023.2184262

Jacob, S., & Furgerson, S. (2012). Writing interview protocols andcConducting interviews: tips for students new to the field of qualitative research. *The Qualitative Report*, *17*(42), 1-10. https://doi.org/10.46743/2160-3715/2012.1718

Jeon, J., Lee, S., & Choi, S. (2023). A systematic review of research on speech-recognition chatbots for language learning: Implications for future directions in the era of large language models. *Interactive Learning Environments*, 1–19. https://doi.org/10.1080/10494820.2023.2204343

Kuhail, M. A., Alturki, N., Alramlawi, S., & Alhejori, K. (2023). Interacting with educational chatbots: A systematic review. *Education and Information Technologies*, *28*(1), 973–1018. https://doi.org/10.1007/s10639-022-11177-3

Li, B., Kou, X., & Bonk, C. J. (2023). Embracing the disrupted language teaching and learning field: Analyzing YouTube content creation related to ChatGPT. *Languages*, *8*(3), 197. https://doi.org/10.3390/languages8030197

Li, M., & Xu, H. (2020). AI-driven language apps and their impact on traditional language learning methods. *Journal of Computer Assisted Learning*, *36*(4), 561–574.

Malik, A. R., Pratiwi, Y., Andajani, K., Numertayasa, I. W., Suharti, S., Darwis, A., & Marzuki. (2023). Exploring artificial intelligence in academic essay: Higher education student's perspective. *International Journal of Educational Research Open*, *5*, 100296. https://doi.org/10.1016/j.ijedro.2023.100296

Morosanu, G.-A., Rata, L.-A., & Geru, M. (2023). Outcomes of large language models and artificial intelligence in education. *Didactica Danubiensis*, *3*(1), 30–52.

Nah, F. F.-H., Zheng, R., Cai, J., Siau, K., & Chen, L. (2023). Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration. *Journal of Information Technology Case and Application Research*, *25*(3), 277–304. https://doi.org/10.1080/15228053.2023.2233814

Ranalli, J. (2021). L2 student engagement with automated feedback on writing: Potential for learning and issues of trust. *Journal of Second Language Writing*, *52*, 100816. https://doi.org/10.1016/j.jslw.2021.100816

Rasul, T., Nair, S., Kalendra, D., Robin, M., Santini, F. de O., Ladeira, W. J., Sun, M., Day, I., Rather, R. A., & Heathcote, L. (2023). The role of ChatGPT in higher education: Benefits, challenges, and future research directions. *Journal of Applied Learning and Teaching*, *6*(1), 41–56. https://doi.org/10.37074/jalt.2023.6.1.29

Shaikh, S., Yayilgan, S. Y., Klimova, B., & Pikhart, M. (2023). Assessing the usability of ChatGPT for formal English language learning. *European Journal of Investigation in Health, Psychology and Education*, *13*(9), 1937–1960. https://doi.org/10.3390/ejihpe13090140

Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology*, *14*(December), 1–14. https://doi.org/10.3389/fpsyg.2023.1260843

Strzelecki, A. (2023). To use or not to use ChatGPT in higher education? A study of students' acceptance and use of technology. *Interactive Learning Environments*, 1–14. https://doi.org/10.1080/10494820.2023.2209881

Trauth, E., & Connolly, R. (2021). Investigating the nature of change in factors affecting gender equity in the IT sector: A longitudinal study of women in Ireland. *MIS Quarterly*, *45*(4), 2055–2100. https://doi.org/10.25300/MISQ/2022/15964

Waluyo, B. (2024). Technology-fused English teaching and learning in higher education: From individual differences to being different individuals. *LEARN Journal: Language Education and Acquisition Research Network*, *17*(2), 42-47. https://so04.tci-thaijo.org/index.php/LEARN/article/view/274071/184852

Waluyo, B., & Kusumastuti, S. (2024). Generative AI in student English learning in Thai higher education: More engagement, better outcomes?. *Social Sciences & Humanities Open*, *10*, 101146. https://doi.org/10.1016/j.ssaho.2024.101146

Xia, Q., Chiu, T. K. F., & Chai, C. S. (2023). The moderating effects of gender and need satisfaction on self-regulated learning through Artificial Intelligence (AI). *Education and Information Technologies*, *28*(7), 8691–8713. https://doi.org/10.1007/s10639-022-11547-x

Yilmaz, F. G. K., Yilmaz, R., & Ceylan, M. (2023). Generative artificial intelligence acceptance scale: A validity and reliability study. *International Journal of Human–Computer Interaction*, 1–13. https://doi.org/10.1080/10447318.2023.2288730

Zaim, M., Arsyad, S., Waluyo, B., Ardi, H., Al Hafizh, M., Zakiyah, M., Syafitri, W., Nusi, A., & Hardiah, M. (2024). AI-powered EFL pedagogy: Integrating generative AI into university teaching preparation through UTAUT and activity theory. *Computers and Education: Artificial Intelligence*, 100335. https://doi.org/10.1016/j.caeai.2024.100335

Zhang, C., Schießl, J., Plößl, L., Hofmann, F., & Gläser-Zikuda, M. (2023). Acceptance of artificial intelligence among pre-service teachers: A multigroup analysis. *International Journal of Educational Technology in Higher Education*, *20*(1), 49. https://doi.org/10.1186/s41239-023-00420-7

Zhang, Z., & Huang, X. (2024). The impact of chatbots based on large language models on second language vocabulary acquisition. *Heliyon*, *10*(3), e25370. https://doi.org/10.1016/j.heliyon.2024.e25370