Contemporary Educational Technology, 2025, **17**(4), ep619
ISSN: 1309-517X (Online)

Research Article

Entangled cognition in EFL education: The role of generative AI

Liwei Hsu [1*]
0000-0002-4901-744X

[1] Graduate Institute of Hospitality Management, National Kaohsiung University of Hospitality and Tourism, Kaohsiung City, TAIWAN
[*] Corresponding author: liweihsu@mail.nkuht.edu.tw

Citation: Hsu, L. (2025). Entangled cognition in EFL education: The role of generative AI. *Contemporary Educational Technology, 17*(4), ep619. https://doi.org/10.30935/cedtech/17621

ARTICLE INFO

Received: 17 Mar 2025

Accepted: 29 Aug 2025

ABSTRACT

This study examines the use of generative artificial intelligence, i.e., ChatGPT, in English as a foreign language (EFL) learning, emphasizing the mediating role of entangled cognition and the effects of the learning outcomes of the tourism students. The research was designed to a quasi-experiment which included 96 participants (48 in an experimental group and 48 in a control group) who were sampled based on convenience to the Spring 2024 semester in one university in southern Taiwan. The "custom virtual language course" experimental group used ChatGPT for personalized language practice and culture learning, control group received traditional learning. A questionnaire package, including the cognitive technology use questionnaire (CTUQ), extended mind scale (EMS), distributed cognition questionnaire (DCQ), metacognitive awareness inventory (MAI), and TOEIC pre- and post-tests was administered to collect the data. The difference-in-differences design was adopted and observed a significant treatment effect such that the treatment group had an average increase in mean scores of 37.98 (standard deviation [SD] = 7.80) compared to 19.62 (SD = 7.80) for the control group and, therefore, an average treatment effect of 21.38 (95% confidence interval [18.74, 24. 01]). Findings suggest that ChatGPT promotes cognitive offloading, distributed cognition, and metacognitive awareness (CTUQ mean [M] = 3.701, EMS M = 3.421, DCQ M = 3.721, MAI M = 3.551), and the development of collaborative learning and cultural competence. These results reveal ChatGPT's potential to reform EFL education, but they also indicate the necessity to mitigate the risks associated with ethical quandaries and over-dependence. Future studies need to create specific scales that can be used for entangled cognition and examine the long-term effects on cognition.

Keywords: ChatGPT, difference-in-differences approach, EFL education, entangled cognition, generative artificial intelligence

INTRODUCTION

Recent breakthroughs in generative artificial intelligence (GenAI) technologies, such as ChatGPT by OpenAI, have revolutionized many professional areas including English as a foreign language (EFL) teaching and learning practices (Huang & Mizamoto 2024; Yang & Li 2024). GenAI's potential in offering adaptive instructional personalization and individualized feedback has the promise to transform EFL learning as it facilitates self-regulated domain knowledge acquisition among different users (Zadoroznyy & Lai 2024). ChatGPT is particularly powerful for natural language processing (NLP) and can be used toward human-like conversation, human-question-answering, and content generation–bridging the gap between theoretical learning through factual knowledge to practical competency in EFL settings. Nonetheless, the successful incorporation of GenAI into education rests on one critical group–students and educators–whose artificial intelligence (AI) literacy (knowledge), awareness and attitudes hold sway over its adoption (Asio, 2024; Asio & Gadia, 2024; Asio & Soriano, 2025).

How students engage with GenAI tools that support functions such as ChatGPT is not only a matter of their AI literacy but also reflects attitudes. Recent studies show that student perceptions of AI are often positive with respect to the instantaneous feedback it can offer and its potential for providing more opportunities to practice languages, but many students lack skills required in order not be swayed by content generated through AIs (Nguyen et al., 2024). In the same way, AI awareness and confidence appear to influence how faculty leverage their role in bringing AI into curricula. Research indicates that educators with stronger AI literacy are inclined to adopt GenAI for educational practices, though there was a considerable worry about academic integrity and excessive dependency as well (Al-Zahrani, 2024). These relationships between stakeholders highlight the necessity of specific training to encourage successful AI integration in EFL teaching.

The notion of entangled cognition, which derives from embodied and extended cognitions (Clark & Chalmers, 1998), is one way of making sense of AI's cognitive significance. Entangled cognition is a hypothesis that cognitive processing takes place in the distributed platform of mind, body and world with AI as an evolving co-orchester (Morais 2023). Although there is a body of research on AI´s roles in lessening cognitive load and optimizing learning efficacy (Chen & Chang, 2024; Feng, 2024), little do we know about how GenAI, ChatGPT in this case establishes entangled cognition for EFL teaching and learning particularly within culturally complex settings such as tourism degree programs.

As for the research gap in this field, the main research in the literature tends to emphasize GenAI's technical functions (Hsu, 2025) or more generalized use in education rather than its potential effect on cognitive engagement within EFL settings. The present study fills this gap by examining the ways in which ChatGPT promotes entangled cognition among EFL tourism students, thereby developing their linguistic and cultural competences. In contrast to previous work that has focused on feedback loop structures or academic writing (Gayed et al., 2022), this research contributes a novel analysis of ChatGPT's interaction via EFL and presents AI-Human cognitive partnerships with fresh insights. Accordingly, this study proposed the following research questions (RQs):

RQ1: How do EFL students perceive ChatGPT's influence on their cognitive engagement with learning content?

RQ2: What are the learning outcomes associated with ChatGPT integration in EFL education?

LITERATURE REVIEW

GenAI and EFL Learning

EFL learning often imposes high cognitive demands due to the complexity of new vocabulary, grammar, and cultural contexts. Recent studies have investigated AI's potential to enhance language acquisition and instructional efficiency. For example, AI-powered feedback systems provide immediate, personalized feedback, promoting self-regulated learning and improving proficiency (Ding & Zou, 2024; Escalante et al., 2023; Wang, 2024). Similarly, AI-driven conversational agents foster learner engagement and enhance oral communication skills through interactive dialogue (Labadze et al., 2023; Shahzad et al., 2024).

GenAI tools further reduce cognitive load by offering personalized feedback and interactive exercises (Chen & Chang, 2024; Feng, 2024). According to Sweller's (1998) cognitive load theory, minimizing extraneous cognitive load enhances learning efficiency. Tools such as AI writing assistants and chatbots enable learners to focus on meaningful language production rather than formulating responses from scratch. Studies confirm that these tools lower extraneous load, allowing learners to engage in deeper cognitive processes essential for language acquisition. Additionally, AI-driven scaffolding supports metacognitive awareness, enhancing learners' self-regulation and progress monitoring (Yang & Xia, 2023). However, the influence of external factors, including GenAI integration, on EFL learners' cognitive mechanisms remains unclear and requires further investigation (Szabó & Szoke, 2024; Yan et al., 2024).

Evolution of AI and Cognitive Science

AI and cognitive science mutually influence each other, driving advancements in understanding human cognition and developing intelligent systems (Chen & Yadollahpour, 2024; Shanmugasundaram & Tamilarasu, 2023; Stevenson et al., 2024). Cognitive science–a multidisciplinary field integrating psychology, neuroscience, linguistics, and computer science–laid the groundwork for AI by examining human intelligence. Early AI models, such as symbolic reasoning systems, were inspired by the cognitive architecture of the human mind, focusing on rule-based logic and problem-solving (Bhuyan et al., 2024; Piantadosi, 2021). However, the limitations of these models, particularly their struggles with ambiguity and adaptability, led researchers to embrace connectionist approaches like neural networks, which mirror the brain's distributed processing mechanisms (Rumelhart et al., 1986). Neural networks have since outperformed symbolic models in tasks requiring pattern recognition, such as speech and image processing (LeCun et al., 2015). Cognitive models have also informed the development of NLP tools by incorporating principles such as semantic memory and contextual learning, enhancing machine learning algorithms' ability to interpret complex linguistic inputs.

Despite these advances, challenges persist at the intersection of AI and cognitive science. AI systems, while effective in specialized tasks, struggle with generalization and transfer learning, which humans perform effortlessly (Alzubaidi et al., 2021; Atchley et al., 2024). Additionally, the "black-box" nature of many AI models raises concerns about transparency and interpretability, essential for understanding cognitive processes (Lipton, 2018). Some researchers argue that current AI models, despite their capabilities, remain far from achieving general intelligence, lacking self-awareness, abstract reasoning, and emotional understanding–core aspects of human cognition (Lake et al., 2017; Tariq et al., 2022). While the interplay between AI and cognitive science has spurred transformative innovations, it also underscores the need for deeper interdisciplinary collaboration to address unresolved challenges and promote responsible AI development.

Entangled Cognition in the AI Era

The concept of entangled cognition, building on theories of embodied and extended cognition, has gained significant traction in this AI era (Manzotti, 2019). This framework proposes that cognition is distributed across the brain, body, and environment, with external tools like AI playing an integral role in shaping human cognition processes. Unlike traditional cognitive models, where the mind is a self-contained processor (Mella, 2020), entangled cognition emphasizes the dynamic interaction between humans and their environments (Vallée-Tourangeau & Vallée-Tourangeau, 2017). As AI systems integrate further with human activities, they form cognitive assemblages that extend cognitive abilities and transform the very nature of cognition (Chen & Yadollahpour, 2024; Zhai et al., 2024). In performing tasks such as memory recall, decision-making, and reasoning, AI permits a new form of distributed intelligence where humans and machines co-constitute cognitive processes (Kejriwal et al., 2024).

In education, for instance, GenAI and adaptive learning platforms act as cognitive scaffolds, helping learners extend their capabilities and personalize their learning experiences (Shen & Wang, 2024). These tools facilitate an intertwined relationship between human cognition and AI, raising opportunities for enhanced learning, and increasing dependency and cognitive autonomy erosion concerns (Shanmugasundaram & Tamilarasu, 2023). Moreover, neuroscience research suggests that frequent interactions with AI and other digital tools may influence neuroplasticity, reshaping the neural architecture of cognition itself (Jain, 2023; Savage, 2019). While entangled cognition provides a valuable framework for understanding the cognitive implications of AI, further research is needed to fully explore how this entanglement impacts human autonomy and agency in the present, evolving AI-driven landscape (Mogi, 2024).

As no scale has yet been developed to measure and rigorously assess entangled cognition, existing instruments such as the cognitive technology use questionnaire (CTUQ), extended mind scale (EMS), distributed cognition questionnaire (DCQ), and metacognitive awareness inventory (MAI) may be used. These instruments are designed to capture the multifaceted interactions between internal cognitive processes and external elements. The CTUQ assesses how individuals integrate technology into their cognitive activities, reflecting the entanglement of human cognition with technological aids (O'Hara et al., 2002; Vongkulluksn et al., 2022). The EMS, rooted in extended cognition theory, evaluates the extent to which external objects and environments extend cognitive processes, and aligns with entangled cognition principles (Clark & Chalmers, 1998; Duus et al., 2018). The DCQ examines how cognitive tasks are distributed across individuals, artifacts, and environments (Hutchins, 1995), given that distributed cognition demonstrates the interconnected nature of cognitive activities (Vasiliou et al., 2015). The MAI, primarily focuses on internal cognitive strategies, but can be adapted to consider the influence of external tools and environments on metacognitive processes (Schraw & Dennison, 1994). Wilson et al. (2020) adopted MAI and explored how distributed cognition involves interactions between individuals and artifacts. They found that tools and environments play a significant role in cognitive processes. These questionnaires provide a comprehensive framework for assessing the dynamic interplay between human cognition and external elements and offer robust methods for measuring entangled cognition.

METHOD

Study Design

This study utilized a quasi-experimental design to examine the influence of GenAI on students' cognitive processes and learning outcomes, with specific focus on entangled cognition. A quasi-experimental design was selected because it offers a practical approach for investigating cause-and-effect relationships in real-world educational settings where the random assignment of participants to experimental and control groups is unfeasible. This enables researchers to compare naturally-occurring groups in educational environments and enhances ecological validity while providing insights into potential causal effects.

In this study, ChatGPT played a role in the experimental group by enhancing both language learning and content knowledge about business operations, such as how to manage employees from various cultural backgrounds. It provided personalized language practice through conversation simulations and instant feedback, helping learners improve their speaking and listening skills. Additionally, ChatGPT offered adaptive learning paths, customized content, and immersive virtual cultural tours, enriching the participants' learning experiences. Overall, ChatGPT created a dynamic, interactive, and supportive learning environment that fostered both linguistic proficiency and cultural awareness. To ensure that the experimental group contained fluent ChatGPT users, they were taught how to craft effective prompts to elicit detailed and contextually appropriate responses from the system, in the first week of the Spring Semester of 2024.

The primary instrument for data collection was a composite questionnaire designed to assess various facets of cognitive engagement with technology, including metacognitive awareness, distributed cognition, and external tool use. This multi-dimensional tool allowed for a nuanced analysis of how students interact with AI tools in educational contexts, capturing both the cognitive processes that occur within learners and those extended or distributed through technology. This methodology ensured the robustness of the study, despite the absence of random assignment; it balanced internal validity and real-world classroom settings.

As for the ethical concern, this study was conducted in accordance with the ethical principles for research involving human subjects and was reviewed and approved by the Institutional Review Board of National Cheng Kung University, under the approval code NCKU-HREC-E-112-690-2. All participants provided their informed consent prior to their involvement in the research.

Participants

To properly address the RQs, the quasi-experimental design was adopted, which is well-suited for educational settings where random assignment is often impractical due to pre-existing class structures or ethical considerations (Cook & Campbell, 1979). This approach allows for the comparison of naturally occurring groups, enhancing ecological validity while exploring causal relationships in real-world contexts (Shadish et al., 2002). The participants were selected using purposive sampling to ensure that both groups had similar demographic and academic backgrounds, thereby minimizing pre-existing differences. The sample size was determined using G*Power software, with parameters set to a medium effect size (d = 0.6), alpha level of .05, and statistical power of .80. These parameters were chosen to ensure sufficient power to detect meaningful differences between two groups and enhance the reliability and validity of the findings. The suggested minimum number was 72 (36 for each group).

Accordingly, two classes of students (n = 96, mean [M_{age}] = 20.3, standard deviation [SD] = 1.8, female = 54, male = 42) from the business-related departments of two universities in Taiwan were invited, and taught by the same instructor. The EFL course was designed to improve students' English proficiency. The syllabi of these two classes were identical, and the weekly course hours were the same at 2 hours per week (see **Table 1**). This semester-long experiment, spanning 18 weeks, was conducted from February to June 2024. However, the participants' academic performance in their EFL course was tracked from September 2023 to facilitate a pre-and post-intervention comparison of their performance. All the participants were briefed on the nature of the research during the first week of the semester (see **Figure 1** for the research procedure).

Table 1. Comparison of intervention and traditional classes

Aspect	Experimental Group (ChatGPT)	Control Group (Traditional)
Teaching procedures	ChatGPT-assisted activities: brainstorming, research, simulated dialogues; supplemented with traditional methods.	Textbook-based exercises, group discussions, and teacher-led lectures.
Instructors' role	Facilitate ChatGPT use, guide prompt creation, monitor interactions, provide feedback.	Deliver lectures, facilitate discussions, provide direct feedback on exercises.
Learning outcome measurement	TOEIC pre- and post-tests (listening and reading, scored 0–100); composite questionnaire (CTUQ, EMS, DCQ, MAI).	TOEIC pre- and post-tests (listening and reading, scored 0–100).

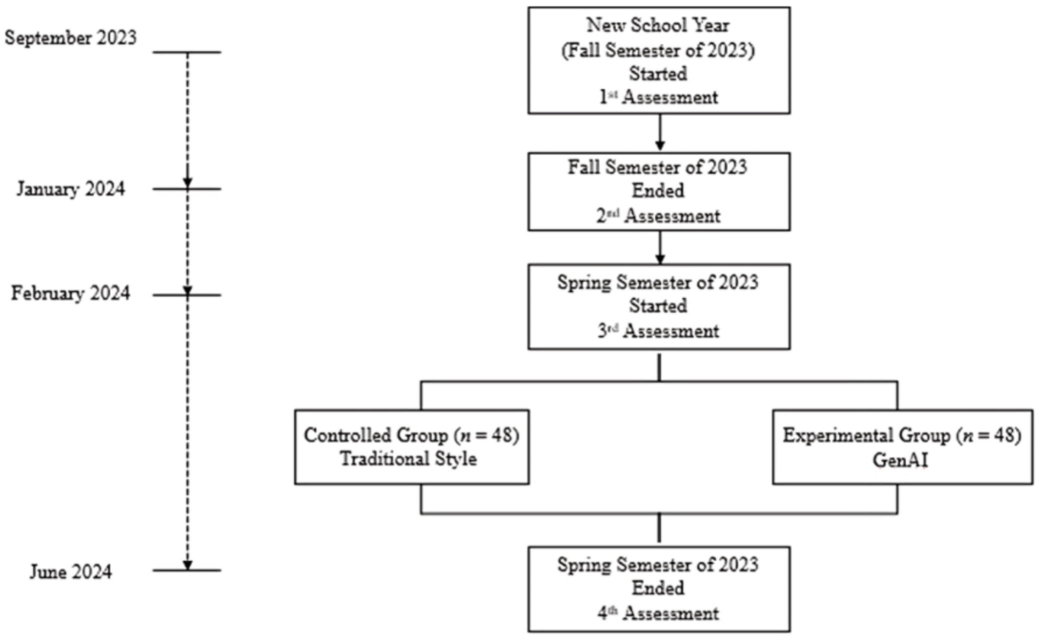

Figure 1. Research procedure (Source: Created by the author)

They were advised that they could withdraw from the experiment without penalty, and that compensation for completion was approximately USD 30.

Measurement

As mentioned above, a specific "entangled cognition" questionnaire does not exist; therefore, the concept must be operationalized through adapted measures from related fields. The measurement used by this study includes items from established scales adapted to measure aspects of entangled cognition, such as the CTUQ (O'Hara et al., 2002), EMS (Clark & Chalmers, 1998), DCQ (Hutchins, 1995), and MAI (Schraw & Dennison, 1994). These scales collectively assess participants' reliance on technology for cognitive tasks, the extent of integration of AI tools with learning, and their overall metacognitive awareness (refer to the Appendix for details on the questionnaire items).

The questionnaire was developed through an extensive review of relevant literature, complemented by feedback from subject matter experts to ensure both theoretical rigor and practical relevance. Internal consistency was assessed using Cronbach's alpha coefficient to evaluate the reliability of the instrument. The analysis yielded a Cronbach's alpha value of 0.74 (against the acceptance threshold of 0.70), indicating a high degree of reliability. This suggests that the questionnaire items were sufficiently interrelated to measure the intended constructs consistently.

A principal component analysis was conducted to explore its underlying factor structure and establish the validity of the instrument. The results confirmed the appropriateness of the proposed factor structure, with all item loadings demonstrating adequate construct validity by exceeding the recommended threshold of 0.50 (Hair et al., 2014). These findings suggest that the items were well-aligned with their respective latent constructs, ensuring that the questionnaire could effectively capture the dimensions it was designed to measure. This assessment was administered in the last week of the semester, and only the experimental group completed this survey.

During the intervention, the experimental group engaged with ChatGPT alongside traditional instruction. Specific applications of ChatGPT included: brainstorming tourism marketing campaigns in various cultural contexts and receiving feedback on ideas; using ChatGPT to research future sustainable tourism in different cultures and summarizing critical findings; and engaging in simulated dialogues with ChatGPT to practice customer service skills relevant to tourism contexts. To measure the participants' learning outcomes, we employed questions comprising TOEIC listening and reading comprehension sections. Each student was graded on a scale of 100 points. Using TOEIC questions to assess EFL learners' learning outcomes has several benefits (Powers & Powers, 2015): it provides a standardized and reliable measure of English proficiency, ensuring consistent and comparable results. It comprehensively assesses various language skills, that align with real-world English usage, which is particularly relevant for practical applications. Global recognition of the test adds value to the scores, making them valuable learner credentials (Powers & Powers, 2015). This evaluation was administered in the first week as a pre-test and in the last week as a post-test.

Data Analysis

Prior to conducting the primary difference-in-differences analysis, comprehensive pretest and posttest analyses were performed to ensure data quality and analytical assumptions were met. For the pretest analysis, descriptive statistics were calculated for both experimental and control groups to examine baseline characteristics. Normality of the pretest scores was assessed using the Shapiro-Wilk test ($W_{experimental}$ = 0.976, p = .341 and $W_{Control}$ = 0.982, p = .673). Independent samples t-tests were conducted to verify baseline equivalence between groups (t(94) = –0.39, p = .696, Cohen's d = 0.08), ensuring no systematic differences existed prior to the intervention. This baseline comparison was crucial for establishing the validity of our quasi-experimental design.

For posttest analysis, similar descriptive analyses were performed on posttest data, including measures of central tendency, variability, and distribution shape. Post-test score distributions were examined for outliers and normality to ensure appropriate statistical modeling. Specifically, no extreme outliers detected using the inter-quartile range (IQR) method (Q1 – 1.5 × IQR to Q3 + 1.5 × IQR) was found in the experimental group while one mild outlier identified (score = 42.5) in the control group; however, this case was retained as it represented a valid data point within the expected range Any extreme values were investigated for data entry errors or other anomalies. Moreover, Skewness and kurtosis of these two groups confirmed the normalized data distribution (skewness = –0.21 and kurtosis = –0.45 for the experimental group while skewness = 0.18, kurtosis = –0.33 for the control group).

The difference-in-differences (DiD) approach is the main statistical analysis method being performed to address the proposed RQs. The approach holds several advantages over techniques such as ANCOVA, particularly in quasi-experimental settings where random assignment is not feasible. DiD effectively controls unobserved confounders that vary over time, offers robustness by not assuming parallel trends in the absence of treatment, and better captures dynamic intervention effects through longitudinal data. This method provides clearer causal interpretations by comparing changes over time between treatment and control groups, addressing limitations in ANCOVA related to time-varying unobserved heterogeneity and the reliance on cross-sectional data (Angrist & Pischke, 2009; Bertrand et al., 2004; Card, 1999; Imbens & Wooldridge, 2009; Wooldridge, 2010).

The DiD analysis can be represented with the following equation: $\Delta Y_{it} = \alpha + \delta D_i + \gamma Post_t + \tau(D_i \times Post_t) + \epsilon_{it}$, where ΔY_{it} refers to change in the learning outcome for individual (*i*) at time (*t*), D_i is a binary indicator where 1 = experimental group, 0 = control group, $Post_t$ represents a binary indicator where 1 = post-intervention period, 0 = pre-intervention period, $D_i \times Post_t$ is the interaction term indicating if the individual is in the experimental group during the post-intervention period.

As for the coefficients and their interpretation: α is the intercept, representing the average outcome for the control group in the pre-intervention period, δ captures the difference in baseline outcomes between the experimental and control groups before the intervention, γ captures the general time effect, representing how outcomes change over time for both groups, τ is the DiD estimator, which measures the *additional*

Table 2. Descriptive statistics of tourism students' entangled cognition when using ChatGPT for learning

	Mean	95% confidence interval mean		Standard deviation	Minimum	Maximum
		Upper	Lower			
CTUQ	3.701	3.877	3.526	0.604	2.330	4.670
EMS	3.421	3.616	3.226	0.672	1.600	5.000
DCQ	3.721	3.909	3.533	0.647	2.000	5.000
MAI	3.551	3.732	3.371	0.622	2.440	4.560

treatment effect of the intervention (e.g., using ChatGPT) on the experimental group. This is the key parameter of interest, as it reflects how much the intervention improved learning outcomes in the experimental group compared to the control group. ϵ_{it} is the error term, capturing unexplained variation.

Descriptive statistics and questionnaire results were analyzed to assess students' perceptions of ChatGPT. All analyses were conducted using JASP (version 0.95.0.0), a widely used statistical software for educational research, ensuring accurate and reproducible results based on R (R Core Team, 2023). No manual computations were performed, and all statistical outputs were verified for accuracy.

RESULTS

The findings offer valuable insights into EFL learners' perceptions of GenAI tools, such as ChatGPT, in enhancing learning outcomes, particularly in fostering global cultural awareness. To address RQ1, four validated instruments–CTUQ, EMS, DCQ, and MAI–were employed. All constructs scored above the midpoint on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree), indicating generally positive attitudes toward ChatGPT as a learning tool.

The CTUQ results (M = 3.701) reveal that learners recognized the cognitive benefits of using ChatGPT, viewing it as a helpful resource for acquiring and organizing business English knowledge. High scores reflect learners' confidence in ChatGPT's ability to retrieve relevant information and facilitate knowledge structuring. The EMS (M = 3.421) highlights ChatGPT's role in cognitive offloading, where learners collaborate with the tool to externalize mental tasks such as idea generation, summarization, and cultural translation. These findings suggest that ChatGPT was not merely an information source but actively supported learners' reflective and problem-solving efforts in culturally complex contexts.

The DCQ (M = 3.721) underscores the importance of distributed cognition, indicating that learners effectively partnered with ChatGPT to manage tasks and co-construct meaning. This suggests growing acceptance of AI-supported collaborative learning, with students comfortably distributing cognitive load between themselves and the tool. The MAI results (M = 3.551) demonstrate learners' ability to regulate their interactions with ChatGPT to optimize learning strategies. Participants reported using the tool to plan, monitor, and evaluate their progress, aligning with research on the importance of metacognitive skills in self-regulated learning environments.

Overall, the results suggest that students successfully integrated ChatGPT into their learning practices, leveraging the tool to enhance their educational experiences. Detailed results are provided in **Table 2**.

For RQ2, this study employed a DiD approach to assess the impact of AI-driven educational tools on students' cognitive processes and learning outcomes, specifically focusing on entangled cognition. The descriptive statistics indicated that the experimental group (n = 48) had mean pre- and post-test scores of 37.63 (SD = 10.01) and 75.83 (SD = 7.80), respectively. The control group (n = 48) had mean pre- and post-test scores of 38.65 (SD = 14.70) and 55.27 (SD = 7.80), respectively.

While comparing the changes in test scores between the experimental and control groups, the mean difference in pre-and post-test scores for the experimental group was 16.63 (SD = 8.18). For the control group, it was 16.63 (SD = 6.76). The average treatment effect, calculated by the difference between the experimental and control groups, was 21.1.58 (95% confidence interval [24.43, 18.74], p < .001).

According to the DiD analysis (**Table 3**), the group coefficient indicates the difference in the outcome variable between the treatment and control groups before the intervention, which reflects the baseline difference between the groups. No significant differences were observed between the two groups. The time coefficient shows the difference in the outcome variable for the control group between the post- and pre-

Table 3. Results of DiD analysis

	Unstandardized coefficient	Standard error	t	p
Intercept α	38.92	2.56	15.22	< .001
Group δ	−1.29	3.62	−0.36	0.722
Time γ	16.42	3.62	4.54	< .001
Group*time τ	21.79	5.11	4.26	< .001

Note. $p < .001$ indicates statistical significance at the 0.1% level (p-value less than 0.001), which is considered highly significant; Standard significance levels are: *$p < .05$ (5% level), **$p < .01$ (1% level), ***$p < .001$ (0.1% level); The group*time interaction coefficient ($\tau = 21.79$) represents the difference-in-differences estimator, measuring the treatment effect of ChatGPT integration on learning outcomes.

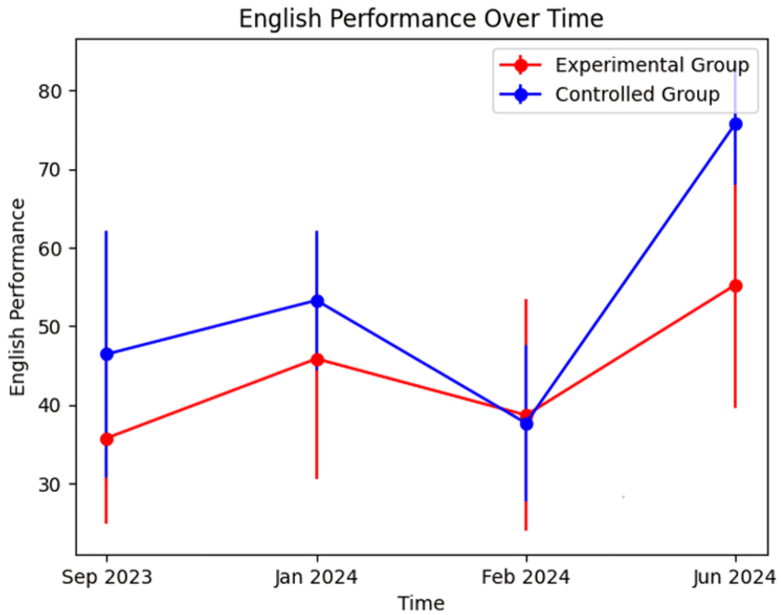

Figure 2. Various time spans of participants' performance (Source: Created by the author)

intervention periods. The results of this study indicate a significant change in the outcome variable over time for the control group ($t = 4.540$, $p < .001$). The interaction coefficient is a critical coefficient in the DiD analysis. It represents the DiD and estimates the treatment effect of the intervention. This coefficient shows how much more (or less) the experimental group changed over time compared to the control group, and a significant result was revealed ($t = 4.261$, $p < .001$). This finding supports the hypothesis that integrating ChatGPT into EFL learning can significantly enhance learners' language skills.

To ensure the robustness of the DiD analysis of this study, several follow-up checks were performed. First, the parallel trends assumption was validated through pre-intervention trend analysis (**Figure 2** presents the pre-intervention trends as parallel; initially indicating that the parallel trends assumption may hold). A placebo test confirmed that no significant effects occurred before the actual intervention, supporting causal inference because of the parallel trends assumption. Finally, attrition checks verified that no participant dropped-out, ruling out selective attrition bias.

DISCUSSION

We reported a quasi-experimental study using difference-in-differences framework to trace the learning trajectory of the ChatGPT group during an 18-week intervention. An experiment was designed and conducted with 96 undergraduate business students from two Taiwanese universities through purposive sampling with 48 participants for each group (experimental and control). We used a composite questionnaire ($\alpha = 0.74$) to collect the data that combined elements from CTUQ, EMS, DCQ, and MAI scales, in addition to TOEIC-based pre- and post-assessments. Statistical analysis of the difference-in-differences was performed in JASP to separate treatment effects from time-invariant confounders. Meanwhile, the lessons took place at two different universities in Taiwan from February to June of 2024 with the same syllabus and teacher for both conditions.

The results of the entangled cognition survey and DiD analysis deliver a clear message that ChatGPT has indeed had an effect as a learning assistant in EFL education. Thus, the survey results indicate that the participants experienced learning benefits from ChatGPT over the semester, which aligns with previous findings on the use of AI tools like ChatGPT in EFL learning contexts (Bin-Hady et al., 2024; Hu & Škultéty, 2024; Karataš et al., 2024; Slamet, 2024). The generally positive view of ChatGPT in cognitive, collaborative, and metacognitive dimensions mirrors broader trends in other empirical evidence available, although the specifics vary.

Similar to previous studies, our findings indicate that EFL learners appreciated the cognitive benefits of ChatGPT, like its facilitation of knowledge organization and retrieval. In the research of Xiao and Zhi (2023), which found that students saw ChatGPT as a conversant with responses in real-time to help with learning activities, especially for the culture and language field. While our participants focused more on leveraging ChatGPT for cultural knowledge, other studies emphasize its role in providing feedback mechanisms for academic writing and grammar enhancement (Gayed et al., 2022). This serves as an indication that while ChatGPT offers versatile cognitive affordances, the focus of its application may depend on the learning context. Furthermore, the findings of this study on cognitive offloading and the concept of the "extended mind" align with those from studies that examined how ChatGPT assists learners by offloading mental tasks. The quality of the grammar and amplitudes of input grammatical error-correction is well-received by learners in EMS, as AI-based collaboration tools seemed trustworthy to enable co-cognition problem-solving processes based on empirical research on writing-assistants (Escalante et al., 2023; Guo et al., 2022). Previous studies have emphasized the use of ChatGPT in individual writing feedback, but our participants pointed out that it has a potential to be implemented even in cases where students are required to learn about business English contexts and this reflects GenAI's flexibility across domains.

Moreover, the high DCQ scores indicate that students worked effectively with ChatGPT, distributing the cognitive load across human-AI systems. This finding aligns with studies that report increased acceptance of AI-driven collaborative learning environments (Huang et al., 2023; Xiao & Zhi, 2023). Despite, there does remain some justifiable apprehension about excessive use of automated responses while working with ChatGPT, a practice that might have the effect of reducing learners' confidence in solving tasks on their own (Kung et al., 2023). However, the results differ slightly as learners are more likely to stay interested in the learning process when ChatGPT is used mainly as a facilitator and not as crutch. Furthermore, with respect to metacognitive awareness revealed from this study, it echoes a point made by Hwang and Chen (2023), that the application of AI technologies such as ChatGPT supports learners in their self-regulation functions, particularly in the process of planning-monitoring-evaluation. However, while prior studies emphasize learners' autonomy when interacting with AI tools for personalized learning (Hsu, 2023; Xiao & Zhi, 2023), our participants specifically valued ChatGPT's support for cultural exploration. This finding broadens the existing understanding of AI tools, indicating that beyond academic writing, these technologies can serve as valuable facilitators in EFL education.

Additionally, the DiD analysis offers evidence of the effectiveness of integrating ChatGPT into EFL learning environments. Initially, no significant differences were observed between the control and experimental groups, ensuring that both started from a comparable baseline. This strengthens the internal validity of the intervention since any observed improvements can be reliably attributed to the use of ChatGPT, rather than pre-existing disparities in language skills. The control group's significant improvement over time ($t = 4.540$, $p < .001$) aligns with findings by Hwang and Chen (2023), who emphasize that consistent practice fosters learning outcomes even in conventional settings. However, the experimental group showed a larger and statistically significant interaction effect ($t = 4.261$, $p < .001$), highlighting that ChatGPT enhances learners' performance beyond what is achievable through traditional methods alone. This improvement can be attributed to GenAI's ability to facilitate reflection, offload cognitive tasks, and support metacognitive practices–echoing Xiao and Zhi's (2023) findings on GenAI's role in personalized feedback, and Hsu's (2023) insights on collaborative writing.

In a nutshell, our findings are consistent with Kung et al. (2023) suggested a potential effectiveness of ChatGPT in learning contexts as it provides real-time cognitive support and scaffolding for problem-solving processes. Together, these findings point to the transformational rather than merely supplementary nature of GenAI tools such as ChatGPT in engaging with tasks more deeply, fostering learner autonomy, and encouraging more collaborative and metacognitive outcomes. These results are consistent with the recent studies depicting that GenAI solutions are helpful in enhancing learning outcomes via adaptive learning and customized assistance (Grassini, 2023).

Implications

The positive outcomes of using ChatGPT in EFL education have several implications with specific guidance for educators and institutions. For educators, the findings of this study may provide the following practical implications. First, the findings highlight AI's potential to bridge the gap between theoretical knowledge and practical skills, providing a more holistic learning experience. This is particularly important in EFL education for business students, where real-world applications and business operation skills are crucial. Educators should integrate ChatGPT strategically into their curriculum by designing tasks that combine language learning with practical business scenarios, such as cross-cultural communication simulations and industry-specific vocabulary development.

Second, ChatGPT's ability to support metacognitive regulation suggests that AI can play a pivotal role in fostering independent and self-directed learning among students. Educators should explicitly train students in prompt engineering and AI interaction strategies to maximize learning benefits. By helping students understand and manage their cognitive processes through guided AI interactions, educators can enhance students' ability to learn more effectively and efficiently (Aoun, 2017; Molenaar, 2022). Practical implementation includes incorporating reflection activities where students analyze their AI-assisted learning processes and develop personalized AI utilization strategies.

Moreover, educational institutions should develop comprehensive AI literacy programs for both faculty and students. Faculty development programs should focus on pedagogical integration of AI tools, ethical AI use, and assessment strategies that account for AI assistance. Students need training in critical evaluation of AI-generated content, understanding AI limitations, and developing AI-human collaborative skills. Furthermore, institutions should establish clear AI usage policies that balance innovation with academic integrity. This includes creating guidelines for appropriate AI use in assignments, developing new assessment methods that leverage AI capabilities while maintaining learning objectives, and implementing support systems for faculty transitioning to AI-integrated pedagogies.

Technical infrastructure investments are crucial, including reliable internet access, AI tool subscriptions, and technical support systems. Institutions should also create AI ethics committees to oversee responsible implementation and address emerging challenges.

Despite these positive findings, it is essential to address the challenges of integrating AI into education. Ethical considerations, data privacy, and the risk of overreliance on AI remain critical issues that must be addressed to ensure the responsible deployment of these technologies (Eubanks, 2018; Floridi & Cowls, 2022). Institutions should implement regular AI literacy workshops, establish clear data governance policies, and create mechanisms for monitoring student dependency on AI tools. Additionally, ongoing research is necessary to monitor the long-term impacts of AI on human cognition and behavior, particularly in educational contexts, where the implications for learning and development are profound (Luckin et al., 2016). Educators should maintain balance between AI assistance and independent cognitive development through scaffolded learning approaches that gradually reduce AI dependence.

Limitations and Future Research Directions

The main limitation of this study is that the findings are based on a relatively small and specific sample, which limits the generalizability of the results. While the data offer valuable insights, a more extensive and diverse sample would provide a more comprehensive understanding of the impact of ChatGPT on tourism education. Another limitation is that the data relied on self-reported measures from the participants, which can introduce biases such as social desirability or inaccurate self-assessment. The measurement of entangled cognition with AI should be specifically developed for an educational context. Furthermore, objective measures of learning outcomes and cognitive changes will help corroborate the findings of this and other studies.

CONCLUSION

This study successfully addressed both RQs through empirical investigation of ChatGPT integration in EFL education using an entangled cognition framework. For addressing RQ1, findings of this study demonstrate that ChatGPT fundamentally shapes cognitive interactions between students and educational content through multiple mechanisms. The entangled cognition survey revealed that students effectively integrated ChatGPT into their cognitive processes across four key dimensions: cognitive technology use ($M = 3.701$), extended mind functioning ($M = 3.421$), distributed cognition ($M = 3.721$), and metacognitive awareness ($M = 3.551$). Students utilized ChatGPT for cognitive offloading, collaborative meaning-making, and metacognitive regulation, indicating that AI tools become genuine cognitive partners rather than mere information sources. Technology facilitated knowledge organization, cultural exploration, and reflective practices, demonstrating how human-AI cognitive entanglement occurs in authentic learning environments.

As for the RQ2, the difference-in-differences analysis provided robust evidence that these cognitive interactions significantly enhanced learning outcomes. While both groups improved over time, the experimental group using ChatGPT showed significantly greater gains (interaction effect: $t = 4.261$, $p < .001$), with an average treatment effect of 21.58 points. This substantial improvement demonstrates that entangled cognitive processes with AI tools translate into measurable learning benefits beyond traditional instruction methods. Putting together, the integration of GenAI, such as ChatGPT, into EFL education showed significant potential to enhance learning experiences and outcomes through entangled cognitive processes. By leveraging the capabilities of AI while maintaining a focus on metacognitive knowledge and regulation, educators can create more effective and inclusive learning environments for students in the digital age. ChatGPT delivers timely, precise, and contextually appropriate answers to inquiries, offering instant and understandable responses that support both linguistic proficiency and cultural awareness. The ability of ChatGPT to tailor learning experiences is significant (Ali et al., 2024; Bettayeb et al., 2024; Mohebi, 2024); however, a consideration remains ChatGPT's limitation of being "incapable of providing adequate factual and/or real-time information regarding the most recent events" (Skavronskaya et al., 2023, p. 255).

Funding: This article was supported by the National Science and Technology Council, Taiwan [113-2410-H-328 -008].

Ethics declaration: This study was approved by the Institutional Review Board with approval code NCKU-HREC-112-690). Informed consent was obtained from all participants prior to their involvement in the study. All personal data were anonymized and stored securely on password-protected servers accessible only to the research team. Data will be retained for five years in accordance with institutional data retention policies.

Declaration of interest: The author declared no competing interest.

Data availability: Data generated or analyzed during this study are available from the author on request.

REFERENCES

Ali, D., Fatemi, Y., Boskabadi, E., Nikfar, M., Ugwuoke, J., & Ali, H. (2024). ChatGPT in teaching and learning: A systematic review. *Education Sciences, 14*(6), Article 643. https://doi.org/10.3390/educsci14060643

Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data, 8*(1), Article 53. https://doi.org/10.1186/s40537-021-00444-8

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press. https://doi.org/10.1515/9781400829828

Aoun, J. E. (2017). *Robot-proof: Higher education in the age of artificial intelligence*. MIT Press. https://doi.org/10.7551/mitpress/11456.001.0001

Asio, J. M. R. (2024). AI literacy, self-efficacy, and self-competence among college students: Variances and interrelationships among variables. *MOJES: Malaysian Online Journal of Educational Sciences, 12*(3), 44–60. https://doi.org/10.22452/aldad.vol12no3.4

Asio, J. M. R., & Gadia, E. D. (2024). Predictors of student attitudes towards artificial intelligence: Implications and relevance to the higher education institutions. *International Journal of Didactical Studies, 5*(2), Article 27763. https://doi.org/10.33902/ijods.202427763

Asio, J. M. R., & Soriano, I. D. (2025). The state of artificial intelligence (AI) use in higher education institutions (HEIs) in the Philippines. In F. D. Mobo (Ed.), *Impacts of AI on students and teachers in education 5.0* (pp. 523–552). IGI Global Scientific Publishing. https://doi.org/10.4018/979-8-3693-8191-5.ch019

Atchley, P., Pannell, H., Wofford, K., Hopkins, M., & Atchley, R. A. (2024). Human and AI collaboration in the higher education environment: Opportunities and concerns. *Cognitive Research: Principles and Implications, 9*(1), Article 20. https://doi.org/10.1186/s41235-024-00547-9

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics, 119*(1), 249–275. https://doi.org/10.1162/003355304772839588

Bettayeb, A. M., Abu Talib, M., Sobhe Altayasinah, A. Z., & Dakalbab, F. (2024, July). Exploring the impact of ChatGPT: Conversational AI in education. *Frontiers in Education, 9*. https://doi.org/10.3389/feduc.2024.1379796

Bhuyan, B. P., Ramdane-Cherif, A., Tomar, R., & Singh, T. P. (2024). Neuro-symbolic artificial intelligence: A survey. *Neural Computing and Applications, 36*(21), 12809–12844. https://doi.org/10.1007/s00521-024-09960-z

Bin-Hady, W. R. A., Ali, J. K. M., & Al-humari, M. A. (2024). The effect of ChatGPT on EFL students' social and emotional learning. *Journal of Research in Innovative Teaching and Learning, 17*(2), 243–255. https://doi.org/10.1108/JRIT-02-2024-0036

Capraro, V., Lentsch, A., Acemoglu, D., Akgun, S., Akhmedova, A., Bilancini, E., Bonnefon, J. F., Brañas-Garza, P., Butera, L., Douglas, K. M., Everett, J. A. C., Gigerenzer, G., Greenhow, C., Hashimoto, D. A., Holt-Lunstad, J., Jetten, J., Johnson, S., Kunz, W. H., Longoni, C., Lunn, P., … Viale, R. (2024). The impact of generative artificial intelligence on socioeconomic inequalities and policy making. *PNAS Nexus, 3*(6), Article pgae191. https://doi.org/10.1093/pnasnexus/pgae191

Card, D. (1999). The causal effect of education on earnings. *Handbook of Labor Economics, 3*, 1801–1863. https://doi.org/10.1016/S1573-4463(99)03011-4

Chen, C. H., & Chang, C. L. (2024). Effectiveness of AI-assisted game-based learning on science learning outcomes, intrinsic motivation, cognitive load, and learning behavior. *Education and Information Technologies, 29*(14), 18621–18642. https://doi.org/10.1007/s10639-024-12553-x

Chen, Z., & Yadollahpour, A. (2024). A new era in cognitive neuroscience: The tidal wave of artificial intelligence (AI). *BMC Neuroscience, 25*(1), Article 23. https://doi.org/10.1186/s12868-024-00869-w

Clark, A. (2008). *Supersizing the mind: Embodiment, action, and cognitive extension*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195333213.001.0001

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis, 58*(1), 7–19. https://doi.org/10.1093/analys/58.1.7

Ding, L., & Zou, D. (2024). Automated writing evaluation systems: A systematic review of Grammarly, Pigai, and Criterion with a perspective on future directions in the age of generative artificial intelligence. *Education and Information Technologies, 29*(11), 14151–14203. https://doi.org/10.1007/s10639-023-12402-3

Duus, R., Cooray, M., & Page, N. C. (2018). Exploring human-tech hybridity at the intersection of extended cognition and distributed agency: A focus on self-tracking devices. *Frontiers in Psychology, 9*. https://doi.org/10.3389/fpsyg.2018.01432

Escalante, J., Pack, A., & Barrett, A. (2023). AI-generated feedback on writing: Insights into efficacy and ENL student preference. *International Journal of Educational Technology in Higher Education, 20*(1), Article 57. https://doi.org/10.1186/s41239-023-00425-2

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

Feng, L. (2024). Investigating the effects of artificial intelligence-assisted language learning strategies on cognitive load and learning outcomes: A comparative study. *Journal of Educational Computing Research, 62*(8), 1741–1774. https://doi.org/10.1177/07356331241268349

Floridi, L., & Cowls, J. (2022). A unified framework of five principles for AI in society. In S. carta (Ed.), *Machine learning and the city: Applications in architecture and urban design* (pp. 535–545). Wiley. https://doi.org/10.1002/9781119815075.ch45

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, 3*, Article 100055. https://doi.org/10.1016/j.caeai.2022.100055

Grassini, S. (2023). Shaping the future of education: Exploring the potential and consequences of AI and ChatGPT in educational settings. *Education Sciences, 13*(7), Article 692. https://doi.org/10.3390/educsci13070692

Guo, K., Wang, J., & Chu, S. K. W. (2022). Using chatbots to scaffold EFL students' argumentative writing. *Assessing Writing, 54*, Article 100666. https://doi.org/10.1016/j.asw.2022.100666

Hsu, H.-C. (2023). The effect of collaborative prewriting on L2 collaborative writing production and individual L2 writing development. *International Review of Applied Linguistics in Language Teaching, 63*(1), 499–533. https://doi.org/10.1515/iral-2023-0043

Hu, W. C., & Škultéty, R. (2024). Unlocking the learning potential: ChatGPT as a virtual platform for cross-interaction in English language learning. *Engineering Proceedings, 74*(1), Article 59. https://doi.org/10.3390/engproc2024074059

Huang, A. Y. Q., Lu, O. H. T., & Yang, S. J. H. (2023). Effects of artificial intelligence-enabled personalized recommendations on learners' engagement, motivation, and outcomes. *Computers and Education, 194*, Article 104684. https://doi.org/10.1016/j.compedu.2022.104684

Huang, J., & Mizumoto, A. (2024). The effects of generative AI usage in EFL classrooms on the L2 motivational self-system. *Education and Information Technologies, 30*, 6435–6454. https://doi.org/10.1007/s10639-024-13071-6

Hutchins, E. (1995). *Cognition in the wild*. MIT Press. https://doi.org/10.7551/mitpress/1881.001.0001

Hwang, G.-J., & Chen, N.-S. (2023). Editorial position paper: Exploring the potential of generative artificial intelligence in education: Applications, challenges, and future research directions. *Educational Technology and Society, 26*(2).

Imbens, G. W., & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature, 47*(1), 5–86. https://doi.org/10.1257/jel.47.1.5

Jain, V. (2023). How AI could lead to a better understanding of the brain. *Nature, 623*(7986), 247–250. https://doi.org/10.1038/d41586-023-03426-3

Karataş, F., Abedi, F. Y., Ozek Gunyel, F., Karadeniz, D., & Kuzgun, Y. (2024). Incorporating AI in foreign language education: An investigation into ChatGPT's effect on foreign language learners. *Education and Information Technologies, 29*, 19343–19366. https://doi.org/10.1007/s10639-024-12574-6

Kejriwal, M., Santos, H., Mulvehill, A. M., Shen, K., McGuinness, D. L., & Lieberman, H. (2024). Can AI have common sense? Finding out will be key to achieving machine intelligence. *Nature, 634*(8033), 291–294. https://doi.org/10.1038/d41586-024-03262-z

Kung, T. H., Cheatham, M., Medenilla, A., Sillos, C., De Leon, L., Elepaño, C., Madriaga, M., Aggabao, R., Diaz-Candido, G., Maningo, J., & Tseng, V. (2023). Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. *PLoS Digital Health, 2*(2), Article e0000198. https://doi.org/10.1371/journal.pdig.0000198

Labadze, L., Grigolia, M., & Machaidze, L. (2023). Role of AI chatbots in education: Systematic literature review. *International Journal of Educational Technology in Higher Education, 20*(1), Article 56. https://doi.org/10.1186/s41239-023-00426-1

Lake, B. M., Ullman, T. D., Tenenbaum, J. B., & Gershman, S. J. (2017). Building machines that learn and think like people. *Behavioral and Brain Sciences, 40*, Article e253. https://doi.org/10.1017/S0140525X16001837

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436–444. https://doi.org/10.1038/nature14539

Lipton, Z. C. (2018). The mythos of model interpretability. *Communications of the ACM, 61*(10), 36–43. https://doi.org/10.1145/3233231

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. UCL Press.

Manzotti, R. (2019). Embodied AI beyond embodied cognition and enactivism. *Philosophies, 4*(3), Article 39. https://doi.org/10.3390/philosophies4030039

Mella, P. (2020). Bateson's model of the mind and the fundamental conjecture on cognition. In *Constructing reality. Springer briefs in psychology* (pp. 1–20). Springer. https://doi.org/10.1007/978-3-030-44132-6_1

Mogi, K. (2024). Artificial intelligence, human cognition, and conscious supremacy. *Frontiers in Psychology, 15*. https://doi.org/10.3389/fpsyg.2024.1364714

Mohebi, L. (2024). Empowering learners with ChatGPT: Insights from a systematic literature exploration. *Discover Education, 3*(1), Article 36. https://doi.org/10.1007/s44217-024-00120-y

Molenaar, I. (2022). Towards hybrid human-AI learning technologies. *European Journal of Education, 57*(4), 632–645. https://doi.org/10.1111/ejed.12527

Morais, R. (2023). Rethinking intelligence beyond the anthropic: Pervasive intelligence, entangled cognition and the logic of the included middle. In L. S. G. M. da Costa, & M. T. Loisel (Eds.), *Artificial intelligence and human mediation* (pp. 105–116). Atlas Press. https://doi.org/10.22545/2024b/B2

Nguyen, A., Hong, Y., Dang, B., & Huang, X. (2024). Human-AI collaboration patterns in AI-assisted academic writing. *Studies in Higher Education, 49*(5), 847–864. https://doi.org/10.1080/03075079.2024.2323593

O'Hara, K., Perry, M., Sellen, A., & Brown, B. (2002a). Exploring the relationship between mobile phones and document activity during business travel. In B. Brown, N. Green, & R. Harper (Eds.), *Wireless world. Computer supported cooperative work* (pp.180–194). Springer. https://doi.org/10.1007/978-1-4471-0665-4_12

Piantadosi, S. T. (2021). The computational origin of representation. *Minds and Machines, 31*(1), 1–58. https://doi.org/10.1007/s11023-020-09540-9

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature, 323*(6088), 533–536. https://doi.org/10.1038/323533a0

Savage, N. (2019). How AI and neuroscience drive each other forwards. *Nature, 571*(7766), S15–S17. https://doi.org/10.1038/d41586-019-02212-4

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology, 19*(4), 460–475. https://doi.org/10.1006/ceps.1994.1033

Shahzad, M. F., Xu, S., & Javed, I. (2024). ChatGPT awareness, acceptance, and adoption in higher education: The role of trust as a cornerstone. *International Journal of Educational Technology in Higher Education, 21*(1), Article 46. https://doi.org/10.1186/s41239-024-00478-x

Shanmugasundaram, M., & Tamilarasu, A. (2023). The impact of digital technology, social media, and artificial intelligence on cognitive functions: A review. *Frontiers in Cognition, 2*. https://doi.org/10.3389/fcogn.2023.1203077

Shen, Y., & Wang, Y. Y. (2024). Navigating the future of higher education: The transformative role of GenAI. In C. K. Y. Chan, & T. Colloton (Eds.), *Generative AI in higher education: The ChatGPT effect*. Routledge. https://doi.org/10.1007/s10734-024-01275-1

Skavronskaya, L., Hadinejad, A., & Cotterell, D. (2023). Reversing the threat of artificial intelligence to opportunity: A discussion of ChatGPT in tourism education. *Journal of Teaching in Travel and Tourism, 23*(2), 253–258. https://doi.org/10.1080/15313220.2023.2196658

Slamet, J. (2024). Potential of ChatGPT as a digital language learning assistant: EFL teachers' and students' perceptions. *Discover Artificial Intelligence, 4*(1), Article 46. https://doi.org/10.1007/s44163-024-00143-2

Stevenson, N., Innes, R. J., Boag, R. J., Miletić, S., Isherwood, S. J. S., Trutti, A. C., Heathcote, A., & Forstmann, B. U. (2024). Joint modelling of latent cognitive mechanisms shared across decision-making domains. *Computational Brain and Behavior, 7*(1), 1–22. https://doi.org/10.1007/s42113-023-00192-3

Szabó, F., & Szoke, J. (2024). How does generative AI promote autonomy and inclusivity in language teaching? *ELT Journal, 78*(4), 478–488. https://doi.org/10.1093/elt/ccae052

Tariq, S., Iftikhar, A., Chaudhary, P., & Khurshid, K. (2022). Examining some serious challenges and possibility of AI emulating human emotions, consciousness, understanding and 'self'. *Journal of Neurology Philosophy, 1*(1), 55–75. https://doi.org/10.5281/zenodo.6637757

Vallée-Tourangeau, G., & Vallée-Tourangeau, F. (2017). Cognition beyond the classical information processing model: Cognitive interactivity and the systemic thinking model (SysTM). In S. Cowley, & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain* (pp. 133–154). Springer. https://doi.org/10.1007/978-3-319-49115-8_7

Vasiliou, C., Ioannou, A., & Zaphiris, P. (2015). An artifact ecology in a nutshell: A distributed cognition perspective for collaboration and coordination. In J. Abascal, S. Barbosa, M. Fetter, T. Gross, P. Palanque, & M. Winckler (Eds.), *Human-computer interaction–INTERACT 2015. INTERACT 2015. Lecture notes in computer science(), vol 9297* (pp. 55–72). Springer. https://doi.org/10.1007/978-3-319-22668-2_5

Vongkulluksn, V. W., Lu, L., Nelson, M. J., & Xie, K. (2022). Cognitive engagement with technology scale: A validation study. *Educational Technology Research and Development, 70*(2), 419–445. https://doi.org/10.1007/s11423-022-10098-9

Wang, C. (2024). Exploring students' generative AI-assisted writing processes: Perceptions and experiences from native and nonnative English speakers. *Technology, Knowledge and Learning, 30*, 1825–1846. https://doi.org/10.1007/s10758-024-09744-3

Wilson, E., Seifert, C., Durning, S. J., Torre, D., & Daniel, M. (2020). Distributed cognition: Interactions between individuals and artifacts. *Diagnosis, 7*(3), 343–344. https://doi.org/10.1515/dx-2020-0012

Wooldridge, J. M. (2010). *Econometric analysis of cross-section and panel data*. MIT Press.

Xiao, Y., & Zhi, Y. (2023). An exploratory study of EFL learners' use of ChatGPT for language learning tasks: Experience and perceptions. *Languages, 8*(3), Article 212. https://doi.org/10.3390/languages8030212

Yan, Y., Sun, W., & Zhao, X. (2024). Metaphorical conceptualizations of generative artificial intelligence use by Chinese university EFL learners. *Frontiers in Education, 9*. https://doi.org/10.3389/feduc.2024.1430494

Yang, L., & Li, R. (2024). ChatGPT for L2 learning: Current status and implications. *System, 124*, Article 103351. https://doi.org/10.1016/j.system.2024.103351

Yang, Y., & Xia, N. (2023). Enhancing students' metacognition via AI-driven educational support systems. *International Journal of Emerging Technologies in Learning, 18*(24), 133–148. https://doi.org/10.3991/ijet.v18i24.45647

Zadorozhnyy, A., & Lai, W. Y. W. (2024). ChatGPT and L2 written communication: A game-changer or just another tool? *Languages, 9*(1), Article 5. https://doi.org/10.3390/languages9010005

Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: A systematic review. *Smart Learning Environments, 11*(1), Article 28. https://doi.org/10.1186/s40561-024-00316-7

APPENDIX A

The questionnaire was developed through a thorough literature review and input from subject matter experts. To assess the reliability of the questionnaire, internal consistency was evaluated using Cronbach's alpha coefficient. The results indicated a high level of reliability, with a Cronbach's alpha value of 0.74, which exceeds the recommended threshold of 0.7. To examine the validity of the questionnaire, a factor analysis was performed to assess the underlying structure of the instrument. The results supported the proposed factor structure, with items loading onto the expected dimensions, demonstrating adequate construct validity. These findings suggest that this questionnaire is a reliable and valid instrument for measuring the construct of interest. The rigorous approach to questionnaire design and evaluation, including the assessment of reliability and validity, strengthens the confidence in the data collected using this instrument and the conclusions drawn from the research.

Cognitive Technology Use Questionnaire

1. How often do you use ChatGPT to complete the assigned tasks?
2. Do you think using technology help you learn English more effectively?
3. I frequently integrate ChatGPT into my learning process.

The Extended Mind Scale

4. Using ChatGPT significantly improves my ability to remember and organize information.
5. I sometimes find it challenging to manage cognitive tasks without relying on ChatGPT.
6. I adapt my use of ChatGPT to suit my cognitive needs and tasks.
7. I feel more confident in my knowledge when I have access to ChatGPT.
8. I often rely on ChatGPT to enhance my memory.

Distributed Cognition Questionnaire

9. I frequently coordinate tasks with ChatGPT to complete complex projects.
10. Effective communication with ChatGPT is crucial for solving problems.
11. Sharing information and resources with ChatGPT helps me to achieve better outcomes.
12. The physical environment I work in significantly affects my productivity and thinking.
13. The design of the classroom impacts how I process information.
14. I frequently consult ChatGPT for specialized knowledge.
15. I am aware of how ChatGPT influences my thinking and decision-making.
16. I consciously adapt my use of ChatGPT and collaboration methods to improve efficiency.
17. I adjust my strategies based on the availability of ChatGPT and resources it provides.

Metacognitive Awareness Inventory

18. I am aware of the strategies that I use ChatGPT when I am studying.
19. I know what the instructor expects me to use ChatGPT to learn.
20. I know how to organize my learning with ChatGPT effectively.
21. I have a specific method for remembering information through ChatGPT.
22. I know when each strategy I use will be most effective when using ChatGPT.
23. I can choose strategies of using GPT that fit the demands of the task.
24. I pace myself while ChatGPT for learning in order to have enough time.
25. I set specific goals of using ChatGPT before I begin a task.
26. I organize my use of ChatGPT to help me understand and remember them.
27. I underline or highlight important information in the ChatGPT.
28. I ask myself periodically if ChatGPT can meet my goals.
29. I change strategies of using ChatGPT when I don't understand.
30. I summarize what I've learned from GPT after I finish.

Table A1. Questionnaire items and reliability and validity examination

Construct	Item	Factor loading	Composite reliability	Average variance extracted
CTUQ (α = 0.742)	I always use ChatGPT to complete the assigned tasks.	0.829	0.833	0.626
	I think using ChatGPT help me learn English more effectively.	0.828		
	I frequently integrate ChatGPT into my learning process.	0.711		
EMS (α = 0.934)	Using ChatGPT significantly improves my ability to remember and organize information.	0.807	0.933	0.737
	I sometimes find it challenging to manage cognitive tasks without relying on ChatGPT.	0.865		
	I adapt my use of ChatGPT to suit my cognitive needs and tasks.	0.877		
	I feel more confident in my knowledge when I have access to ChatGPT.	0.877		
	I often rely on ChatGPT to enhance my memory.	0.865		
DCQ (α = 0.840)	I frequently coordinate tasks with ChatGPT to complete complex projects.	0.514	0.875	0.590
	Effective communication with ChatGPT is crucial for solving problems.	0.831		
	Sharing information and resources with ChatGPT helps me to achieve better outcomes.	0.831		
	The physical environment I work in significantly affects my productivity and thinking.	0.807		
	I consciously adapt my use of ChatGPT and collaboration methods to improve efficiency.	0.807		
MAI (α = 0.795)	I am aware of the strategies that I use ChatGPT when I am studying.	0.782	0.869	0.508
	I know when each strategy I use will be most effective when using ChatGPT.	0.698		
	I know how to organize my learning with ChatGPT effectively.	0.748		
	I pace myself while ChatGPT for learning in order to have enough time.	0.716		
	I organize my use of ChatGPT to help me understand and remember them.	0.619		
	I summarize what I've learned from GPT after I finish.	0.702		