

Integrating Generative AI and Gamification to Enhance BESTEP Performance: A Case Study of a College English Empowerment Course

Wan-Yu Chang

Department of Education and Learning Technology, National Tsing Hua University, Hsinchu City, Taiwan,
R.O.C

Email: changwy@mx.nthu.edu.tw

Abstract

This study investigates the effects of integrating generative artificial intelligence (AI) and gamification on college students' performance in the BEST Test of English Proficiency (BESTEP). Taking an English Empowerment course at a university in Taiwan as a case study, ChatGPT was incorporated to support speaking practice and self-directed learning, while gamified platforms such as Kahoot, Blooket, and Gimkit were used for vocabulary and grammar reinforcement. A quantitative research approach was adopted, with 37 valid questionnaires collected and analyzed using correlation and regression analyses. The results indicate that both gamification and ChatGPT-supported learning significantly enhance students' learning motivation and self-directed learning readiness, which in turn positively influence BESTEP performance. Notably, ChatGPT demonstrated the strongest effect on self-directed learning ($\beta = 0.89$, $p < .01$). Furthermore, the findings reveal a dual-path learning model characterized by "high-frequency engagement" and "deep empowerment," which effectively improves students' learning confidence and overall performance. The results of this study provide important pedagogical implications for bilingual education in higher education.

Keywords: generative artificial intelligence, gamification, BESTEP, self-directed learning, learning motivation

1. Introduction

With the rapid development of globalization and increasing international competition, enhancing college students' English proficiency has become a key objective in higher education. To strengthen students' English abilities and global competitiveness, Taiwan's Ministry of Education launched the Program on Bilingual Education for Students in College (BEST Program, 2021) and developed the BEST Test of English Proficiency (BESTEP, 2023) as an important assessment of students' listening and reading skills. However, in practical teaching contexts, English Empowerment courses still face several challenges, including low learning motivation, passive learning approaches, limited opportunities for speaking practice, and foreign language anxiety, all of which hinder learning effectiveness.

In recent years, generative artificial intelligence (AI), such as ChatGPT, and gamification tools, such as Kahoot, Blooket, and Gimkit, have been increasingly applied in language education. ChatGPT provides immediate feedback and low-anxiety conversational environments, facilitating speaking practice and self-directed learning. Meanwhile, gamified platforms enhance student engagement and motivation through competition and interactive mechanisms.

However, existing studies have primarily focused on the application of these tools individually, with limited research examining their integrated use in instructional design. Therefore, this study aims to investigate the effects of integrating ChatGPT and gamification into an English Empowerment course on students' learning motivation, self-directed learning readiness, and BESTEP performance.

2. Literature Review

2.1 Applications of Generative AI in Language Learning

With the rapid advancement of generative artificial intelligence, tools such as ChatGPT have been widely adopted in language learning contexts. Previous studies indicate that AI can provide immediate feedback, personalized learning content, and interactive language practice, thereby enhancing students' writing ability, speaking fluency, and learning motivation (Wang & Lee, 2024). In particular, ChatGPT is capable of simulating authentic conversational contexts (Jeon & Lee, 2023), enabling learners to practice language in low-anxiety environments, which in turn increases opportunities for language output and boosts learner confidence (Tai & Chen, 2024).

Generative artificial intelligence (GAI), especially ChatGPT based on large language models (LLMs), has emerged as a transformative tool in language education. Recent studies suggest that the application of GAI in English language teaching can be examined from the following three key dimensions:

2.1.1 Providing a Low-Anxiety and Personalized Speaking Environment

In traditional language learning contexts, students often experience high levels of Foreign Language Classroom Anxiety (FLCA) due to fear of making mistakes or being evaluated by peers (Horwitz, Horwitz, & Cope, 1986), which in turn inhibits their willingness to speak and participate. Such anxiety not only affects learners' language output but may also reduce their learning motivation and confidence. Hsu (2023) pointed out that generative AI tools such as ChatGPT can serve as a "non-threatening" learning environment and act as a non-judgmental interlocutor, allowing learners to practice language without pressure, thereby reducing anxiety and increasing engagement.

Compared to traditional in-class speaking activities, ChatGPT offers a more flexible and personalized learning environment. Learners can adjust the content and difficulty of conversations according to their own proficiency levels and learning needs, and gradually improve their speaking ability through repeated practice. In addition, Wang and Lee (2024) addressed that the immediate responses and continuous interaction provided by AI enable students to learn from their mistakes without the pressure of real-time evaluation from teachers or peers, thereby fostering a safe and supportive learning environment.

2.1.2 Scaffolding Theory and Adaptive Content Generation

A key advantage of generative artificial intelligence (GAI) in language education lies in its ability to generate adaptive learning materials tailored to learners' proficiency levels. This

aligns with Vygotsky's (1978) Zone of Proximal Development (ZPD), which emphasizes learning within the range between current ability and potential development, supported through scaffolding (Wood, Bruner, & Ross, 1976).

In contrast to traditional instruction, where materials are often fixed and inflexible, generative AI enables dynamic adjustment of content difficulty. Large language models (LLMs) can provide personalized and responsive learning experiences, functioning as a form of digital scaffolding that adapts to learner input (Kasneci et al., 2023).

Accordingly, ChatGPT can generate vocabulary, sentence structures, and conversational content aligned with specific proficiency levels (e.g., CEFR A2–B1) and adjust responses in real time (Jeon & Lee, 2023). In this study, prompts were designed to control language level and response length, ensuring alignment with BESTEP requirements while providing continuous scaffolding in both comprehension and production.

Furthermore, AI-driven dynamic scaffolding through prompt engineering allows precise control over content complexity and supports iterative learning (Su & Yang, 2023). Combined with immediate feedback and diverse language input, generative AI facilitates gradual language development, enhances learner confidence, and sustains motivation, particularly for A2–B1 learners.

2.1.3 Immediate Feedback and the Development of Self-Directed Learning

Immediate feedback is a critical mechanism in language learning, as it enhances learning effectiveness and promotes self-regulation. Generative artificial intelligence (GAI) provides real-time feedback that helps learners identify errors, engage in self-correction, and continuously improve their language proficiency, addressing the limitations of delayed feedback in traditional classrooms (Wang & Lee, 2024; O'Neill & Russell, 2024).

In the English Empowerment course, students used ChatGPT to compose dialogue scripts and record spoken interactions, integrating “input processing” (VanPatten, 2004) and “output practice” (Swain, 1985). This interaction enables learners to process input and produce output in real time, forming an “acquisition loop” that strengthens language internalization and application (Tai & Chen, 2024; Wang & Lee, 2024).

Moreover, ChatGPT provides multi-layered feedback, including grammatical correction and enriched linguistic input such as vocabulary and pragmatic expressions. This supports a shift from passive reception to active knowledge construction and enhances learners' self-directed learning abilities (Tai & Chen, 2024; Wang & Lee, 2024).

In sum, through immediate feedback and the integration of input and output processes, generative AI facilitates more efficient, interactive, and self-regulated language learning.

2.2 Gamification Theory and Vocabulary Acquisition

Gamification is not merely the use of games; rather, it refers to the application of game design elements in non-game contexts (Deterding et al., 2011) to stimulate learners' psychological engagement. In English Empowerment courses, gamified instruction incorporates elements

such as points, challenges, and competition to satisfy learners' needs for autonomy and competence (Deci & Ryan, 2000), while transforming otherwise tedious vocabulary review into strategic and engaging learning tasks (Kapp, 2012).

This approach not only enhances student engagement but also strengthens language acquisition through high-frequency interaction and repeated practice (Huang et al., 2020). Gamification improves learning outcomes primarily through the following mechanisms:

2.2.1 Self-Determination Theory and Learning Motivation

According to Self-Determination Theory (SDT) proposed by Deci and Ryan (2000), effective learning motivation arises from the fulfillment of three basic psychological needs: autonomy, competence, and relatedness. In the present study, the real-time leaderboard feature in Kahoot provides visible indicators of achievement, thereby enhancing learners' sense of competence (Sailer et al., 2017). Meanwhile, platforms such as Blooket and Gimkit allow students to manage in-game resources, such as items and virtual currency, which strengthens their sense of autonomy.

This mechanism effectively transforms the otherwise tedious process of vocabulary repetition, as highlighted by Nation (2001), into a strategic and engaging challenge. As a result, learners unconsciously engage in high-frequency language practice while striving to achieve game-related goals, which in turn significantly enhances their learning motivation.

2.2.2 High-Frequency Repetition and Long-Term Retention

Research in language acquisition suggests that new vocabulary requires approximately 5 to 16 meaningful encounters to be transferred into long-term memory (Nation, 2001). Traditional instructional approaches often fail to provide sufficient and engaging opportunities for such repetition. In contrast, the game-loop design of platforms like Gimkit and Blooket enables learners to engage in more than 20 high-frequency interactions within a short period of time.

This mechanism incorporates the retrieval practice effect (Roediger & Karpicke, 2006), requiring learners to rapidly recall language information under time constraints, thereby promoting the development of automaticity in language processing (Reinhardt & Sykes, 2012). The findings of this study further indicate that such high-frequency repetitive practice can effectively translate into improved listening and reading performance in the BESTEP test, while also significantly enhancing long-term vocabulary retention (Huang et al., 2020).

2.2.3 The Role of Digital Gamification in Promoting Self-Directed Learning

According to Self-Regulated Learning (SRL) theory proposed by Barry Zimmerman (2000), effective learning requires learners to engage in self-monitoring, self-reaction, and self-adjustment. In gamified learning environments, real-time feedback provided by digital platforms—such as scores and leaderboards—serves as a critical source of information for learners' self-monitoring processes (Butler & Winne, 1995).

In the present study, it was observed that students engaged in asynchronous practice through Blooket and Gimkit were able to iteratively correct errors and refine their learning strategies

based on performance feedback. This process not only enhanced learning engagement but also significantly increased learners' initiative in planning and regulating their own learning progress. These findings are consistent with the study by Laura-De La Cruz et al. (2023), which confirms that gamification mechanisms can effectively support learners' self-regulation processes and, in turn, strengthen their self-directed learning readiness (Bandura, 1997).

2.2.4 Effectiveness of Gamification in English Language Teaching

Gamification refers to the integration of game elements—such as points, competition, and rewards—into educational contexts to enhance motivation and engagement (Deterding et al., 2011). Prior research has shown that gamification significantly improves English learning outcomes, particularly in vocabulary retention and learner engagement (Huang et al., 2020).

In listening and reading training, platforms such as Kahoot and Blooket incorporate time constraints and immediate feedback, facilitating rapid retrieval of language knowledge and supporting skills required for the BEST Test of English Proficiency (BESTEP). Moreover, gamified learning reduces anxiety and enhances engagement by creating a low-anxiety, high-engagement environment, consistent with the affective filter hypothesis. Interactive and competitive game designs further promote both behavioral and cognitive engagement, encouraging learners to analyze errors and deepen understanding. Additionally, gamification demonstrates cross-skill transfer effects, improving not only vocabulary performance but also learners' background knowledge and confidence.

In summary, gamified learning enhances engagement, reinforces language retention, and promotes active learning, providing a strong foundation for BESTEP listening and reading performance.

3. Research Method

3.1 Research Framework and Hypotheses

This study proposes an integrated model that combines in-class gamified practice with out-of-class AI-supported self-directed learning. In the proposed framework, Gamified Vocabulary Practice (GVP) and ChatGPT-Supported Self-Directed Learning (CSDL) are treated as exogenous variables, while Learning Motivation (LM) and Self-Directed Learning Readiness (SDLR) function as mediating variables. BESTEP Learning Performance (BLP) is specified as the outcome variable.

Based on the literature review, the following six hypotheses are proposed:

H1: Gamified vocabulary practice positively affects learning motivation.

H2: Gamified vocabulary practice positively affects self-directed learning readiness.

H3: ChatGPT-supported self-directed learning positively affects learning motivation.

H4: ChatGPT-supported self-directed learning positively affects self-directed learning readiness.

H5: Learning motivation positively affects BESTEP learning performance.

H6: Self-directed learning readiness positively affects BESTEP learning performance.

3.2 Participants and Context

The participants in this study were undergraduate students enrolled in an English Empowerment course at Chung Hua University in Taiwan ($N \approx 40$). The students' English proficiency ranged from CEFR levels A2 to B1. The course spanned 18 weeks, with two hours of instruction per week, and was designed to prepare students for the BEST Test of English Proficiency (BESTEP) promoted by Taiwan's Ministry of Education.

3.3 Instructional Intervention

The instructional design integrates two core strategies: gamified learning and generative AI-supported learning.

- 1) Gamified Vocabulary and Grammar Practice (GVP)
After each class, students engaged in immediate practice using Kahoot, Bloocket, and Gimkit, where competitive elements and real-time feedback reinforced vocabulary and sentence structures required for BESTEP.
- 2) ChatGPT-Supported Speaking Task (CSDL)
For the midterm, students completed a four-minute English conversation using ChatGPT voice interaction. Prompts were designed at the CEFR A2–B1 level, with responses limited to 40 words, and recordings were uploaded to Moodle for evaluation.
- 3) Final Examination (BLP)
The final assessment was a BESTEP-based paper test, consisting of listening (50%) and reading (50%), serving as the measure of learning performance.

3.4 Instruments and Validation

This study employed a five-point Likert scale to collect data across five constructs: Gamified Vocabulary Practice (GVP), ChatGPT-Supported Learning (CSDL), Learning Motivation (LM), Self-Directed Learning Readiness (SDLR), and BESTEP Learning Performance (BLP). Each construct consisted of four items, while BLP was measured using both final examination scores and self-reported learning outcomes.

To ensure the reliability and validity of the measurement model, several statistical tests were conducted, including Cronbach's alpha ($\alpha > .70$) for internal consistency, average variance extracted ($AVE > .50$) for convergent validity, and composite reliability ($CR > .70$). Discriminant validity was assessed using the Fornell–Larcker criterion.

3.5 Data Collection and Analysis

In Week 1, participants were informed of the study purpose and provided consent. Following a 14-week instructional intervention and a midterm task, final examination and questionnaire data were collected in Week 18. The study integrated gamified learning and AI-supported learning, combining in-class gamified practice with out-of-class ChatGPT-supported self-directed

learning. It was hypothesized that gamified practice and AI-supported learning would enhance learning motivation and self-directed learning readiness, thereby improving BESTEP performance.

A mixed-methods approach was adopted. Quantitative data were analyzed using descriptive statistics and confirmatory factor analysis (CFA) to assess reliability and validity, followed by correlation and multiple regression analyses to test hypotheses H1–H6 and their explanatory power (R^2). Bootstrap analysis was conducted to examine mediating effects. Qualitative data were analyzed through thematic analysis to explore students' learning experiences and perceptions, complementing the quantitative findings.

3.6 Ethical Considerations

All participants took part in the study on a voluntary basis with informed consent. All data were anonymized to ensure confidentiality and to uphold the rigor and ethical standards of academic research.

4. Finding and Discussion

4.1 Descriptive Statistics and Reliability Analysis

A total of 37 valid questionnaires were collected in this study. The results of the descriptive statistical analysis indicated that students gave highly positive evaluations across all instructional intervention dimensions, with mean scores ranging from 4.27 to 4.48. Among these, Gamified Vocabulary Practice (GVP) received the highest mean score ($M = 4.48$, $SD = 0.52$), suggesting that the competitive features of platforms such as Blooket and Gimkit were highly favored by students.

The mean score for ChatGPT-Supported Speaking (CSDL) was 4.36, indicating that AI-based voice interaction tasks were effective in engaging learners at the CEFR A2–B1 levels.

In terms of reliability analysis, the Cronbach's alpha values for all constructs exceeded 0.85, demonstrating excellent internal consistency of the measurement scales.

4.2 Hypothesis Testing

Pearson correlation analysis and multiple regression analysis were conducted to examine the hypothesized relationships (H1–H6) in this study. The results are summarized in Table 1.

Table 1. Summary of Hypothesis Testing Results

Hypothesis	Path	Statistical Result (β / r)	Result
H1	Gamified Practice → Learning Motivation	0.82**	Supported
H2	Gamified Practice → Self-Directed Learning Readiness	0.78**	Supported
H3	ChatGPT Support → Learning Motivation	0.85**	Supported

H4	ChatGPT Support → Self-Directed Learning Readiness	0.89**	Supported
H5	Learning Motivation → BESTEP Learning Performance	0.81**	Supported
H6	Self-Directed Learning Readiness → BESTEP Learning Performance	0.84**	Supported

Note: $p < .01$, indicating statistical significance.

The results indicate that all proposed hypotheses (H1–H6) were supported, demonstrating strong positive relationships among the key variables.

4.3 In-Depth Data Analysis and Discussion

4.3.1 Gamification as a Driver of High-Frequency Engagement (H1, H2, H5)

The results indicate a strong consistency in students' positive responses to items such as "I enjoy practicing vocabulary through Blooket/Gimkit" and "This makes me more motivated to learn English." These findings confirm that key gamification elements—such as competition and strategic gameplay (e.g., the accumulation mechanisms in Gimkit)—effectively reduce the perceived monotony of vocabulary review for BESTEP preparation and successfully transform external incentives into intrinsic learning engagement.

Moreover, the real-time feedback mechanism enables learners to rapidly identify and correct errors during listening and reading practice, thereby enhancing the effectiveness of high-frequency engagement. This continuous interaction not only reinforces vocabulary retention but also strengthens learners' overall participation in the learning process.

4.3.2 ChatGPT as a Low-Anxiety Environment for Self-Directed Learning (H3, H4, H6)

Among all hypothesized relationships, H4 (the effect of ChatGPT on self-directed learning readiness) demonstrated the strongest correlation ($r = 0.89$). This finding suggests that students' sense of learning agency was significantly enhanced when they engaged with AI through Voice Control and prompt-based interactions on topics such as honeymoon travel and smart home living.

As a non-judgmental interlocutor, ChatGPT substantially reduces learners' language anxiety, enabling them to engage in repeated practice and self-monitor their learning progress. This process fosters a supportive and low-anxiety learning environment, allowing students to take greater control of their learning and ultimately achieve the goal of "deep empowerment."

4.3.3 Overall Impact on BESTEP Learning Performance

The analysis indicates that students who actively participated in both gamified learning tasks and AI-supported speaking activities demonstrated significant improvements in listening and reading comprehension in the final BESTEP mock test. Students generally perceived that AI-based tasks enhanced their speaking confidence, while gamified platforms improved their speed and accuracy in vocabulary and grammar retrieval.

Overall, students reported a notable increase in their confidence in preparing for BESTEP, with a mean score of 4.27, reflecting a significant enhancement in their perceived learning performance.

5. Conclusion

This study examined the effectiveness of integrating generative AI and gamification in a college English Empowerment course. The findings support a dual-path model that significantly enhances learning motivation, self-directed learning readiness, and BESTEP performance, offering important implications for bilingual education. ChatGPT provides a low-anxiety, interactive environment that supports sustained language practice and shows the strongest effect on self-directed learning readiness ($\beta = 0.89$, $p < .01$), conceptualized as “deep empowerment.” In contrast, gamified platforms such as Kahoot, Blooket, and Gimkit promote high-frequency interaction and real-time feedback, enhancing engagement and reinforcing foundational knowledge, conceptualized as “high-frequency engagement.” Together, these two pathways integrate language input and output, forming a complete learning cycle that reduces anxiety, increases engagement, and improves learning outcomes for A2–B1 learners. Overall, the integration of generative AI and gamification represents a practical and scalable instructional strategy for enhancing English learning and advancing the BEST Program in higher education.

Limitations. Despite the significant findings, several limitations should be noted. First, the study was conducted in a single university with a relatively small sample size ($N = 37$), which may limit generalizability. Future research should include larger and more diverse samples across institutions and proficiency levels. Second, the intervention lasted only one semester, making it difficult to assess long-term effects on BESTEP performance; longitudinal studies are recommended. Finally, the study employed a single-group design without a control group. Future research should adopt quasi-experimental or experimental designs to better establish causal relationships.

Pedagogical Suggestions. Based on the findings, several pedagogical implications are proposed. First, instructors are encouraged to design scenario-based AI dialogue tasks (e.g., career or real-life contexts) to enhance authenticity and practical language use. Second, rotating multiple gamification platforms (e.g., Kahoot, Blooket, Gimkit) is recommended to sustain engagement and novelty. Finally, adequate technical support should be provided at the initial stage, including guidance on prompt design and voice interaction tools, to reduce barriers and ensure effective implementation.

Theoretical Contribution. This study contributes to the literature on technology-enhanced language learning in three ways. First, it reconceptualizes generative AI as a mechanism of “deep empowerment,” demonstrating how ChatGPT supports learners’ self-directed learning readiness and metacognitive development, thereby extending Self-Regulated Learning (SRL) theory. Second, it reframes gamification as “high-frequency engagement,” showing how game-based mechanisms enhance motivation and repeated language exposure, which are essential for vocabulary retention and language acquisition, grounded in Self-Determination Theory (SDT). Third, this study proposes and validates a dual-path model that integrates generative AI and

gamification. By positioning AI as facilitating output and self-regulation, and gamification as strengthening input and engagement, this model offers a novel framework for understanding how technology supports a complete language learning cycle.

References

- Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman and Company.
- Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, 65(3), 245–281. <https://doi.org/10.3102/00346543065003245>
- Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01
- Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: Defining "gamification". In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments* (pp. 9–15). Association for Computing Machinery. <https://doi.org/10.1145/2181037.2181040>
- Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, 70(2), 125–132. <https://doi.org/10.1111/j.1540-4781.1986.tb05256.x>
- Hsu, L. (2023). ChatGPT as a chatbot for language learning: Benefits and challenges. *Computers and Education: Artificial Intelligence*, 5, 100150. <https://doi.org/10.1016/j.caeai.2023.100150>
- Huang, R., Ritzhaupt, A. D., Sommer, M., Zhu, J., Stephen, A., Diehl, G., & Gordan, J. (2020). The effects of gamification on vocabulary learning: A meta-analysis. *Educational Research Review*, 30, 100326. <https://doi.org/10.1016/j.edurev.2020.100326>
- Jeon, J. H., & Lee, S. (2023). Large language models in foreign language learning: Real-world dialogues and personalized feedback. *Computer Assisted Language Learning*. Advance online publication. <https://doi.org/10.1080/09588221.2023.2265022>
- Jeon, J. H., & Lee, S. (2023). Large language models in foreign language learning: Real-world dialogues and personalized feedback. *Computer Assisted Language Learning*. <https://doi.org/10.1080/09588221.2023.2265022>
- Kapp, K. M. (2012). *The gamification of learning and instruction: Game-based methods and strategies for training and education*. Pfeiffer.
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., ... & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274.
- Language Training and Testing Center. (2023). *The development and validation of the Best Test*

of English Proficiency (BESTEP). <https://www.lttc.ntu.edu.tw/BESTEP.htm>

Laura-De La Cruz, G. A., Gamarra-Chavez, M. A., Vera-Flores, M. A., & Cruzate-Velez, C. V. (2023). Gamification and self-regulated learning in higher education: A systematic review. *Journal of Technology and Science Education*, 13(1), 321–337. <https://doi.org/10.3926/jotse.1895>

Ministry of Education. (2021). *The Program for Bilingual Education for Students in College (BEST Program)*. <https://best.moe.gov.tw/>

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139524759>

O'Neill, E. M., & Russell, A. (2024). Leveraging generative AI for immediate corrective feedback in language learning. *Language Learning & Technology*, 28(1), 1–15.

Reinhardt, J., & Sykes, J. M. (2012). Conceptualizing digital game-based L2 learning and outcomes. *NECTFL Review*, 70, 32–62.

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, 17(3), 249–255. <https://doi.org/10.1111/j.1467-9280.2006.01693.x>

Sailer, M., Hense, J. U., Mayr, S. K., & Mandl, H. (2017). How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction. *Computers in Human Behavior*, 69, 371–380. <https://doi.org/10.1016/j.chb.2016.12.033>

Su, J., & Yang, W. (2023). Unlocking the power of ChatGPT: A framework for leveraging large language models in education. *International Journal of Engineering Business Management*, 15, 1–14. <https://doi.org/10.1177/18479790231167547>

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in Second Language Acquisition* (pp. 235–253). Newbury House.

Tai, K. W. H., & Chen, L. (2024). Enhancing language output through AI chatbots: A study on learner autonomy and confidence in EFL contexts. *System*, 121, 103248. <https://doi.org/10.1016/j.system.2024.103248>

Tai, K. W., & Chen, L. (2024). Enhancing language output through AI chatbots: A study on learner autonomy and confidence in EFL contexts. *System*, 121, 103240. <https://doi.org/10.1016/j.system.2024.103240>

VanPatten, B. (2004). Input processing in adult second language acquisition. In B. VanPatten (Ed.), *Processing instruction: Theory, research, and commentary* (pp. 5–31). Lawrence Erlbaum Associates.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*.

Harvard University Press.

Wang, Y., & Lee, J. S. (2024). The impact of generative AI chatbots on English as a foreign language (EFL) learners' writing and speaking performance: A mixed-methods study. *Journal of Educational Technology & Society*, 27(2), 145–160. [https://doi.org/10.30191/ETS.202404_27\(2\).0010](https://doi.org/10.30191/ETS.202404_27(2).0010)

Wang, Y., & Lee, J. S. (2024). The impact of generative AI chatbots on English as a foreign language (EFL) learners' writing and speaking performance: A mixed-methods study. *System*, 121, 103240.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, 17(2), 89–100.

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). Academic Press. <https://doi.org/10.1016/B978-012109890-3/50031-7>

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>)