Routledge
Taylor & Francis Group

Innovation in Language Learning and Teaching

ISSN: 1750-1229 (Print) 1750-1237 (Online) Journal homepage: www.tandfonline.com/journals/rill20

Enhancing HOTS through role-playing GenAI: a case study in classical literature education

Zhaoyang Xiong, Junjie Gavin Wu, Sumei Wu & Junhua Xian

To cite this article: Zhaoyang Xiong, Junjie Gavin Wu, Sumei Wu & Junhua Xian (28 Jan 2026): Enhancing HOTS through role-playing GenAI: a case study in classical literature education, Innovation in Language Learning and Teaching, DOI: 10.1080/17501229.2026.2619908

To link to this article: https://doi.org/10.1080/17501229.2026.2619908

INNOVATION IN LANGUAGE LEARNING AND TEACHING
https://doi.org/10.1080/17501229.2026.2619908

Routledge
Taylor & Francis Group

REPORT

Enhancing HOTS through role-playing GenAI: a case study in classical literature education

Zhaoyang Xiong [a], Junjie Gavin Wu [a], Sumei Wu [b] and Junhua Xian [a]

[a]Faculty of Applied Sciences, Macao Polytechnic University, Macao, Macao SAR; [b]School of Foreign Languages and Literature, Beijing Normal University, Beijing, People's Republic of China

ABSTRACT

Purpose: This educational innovation introduces a customized role-playing GenAI that simulates characters from texts, operating under the premise that explicit instruction in classical literature education with an emphasis on active, learner-centered learning can cultivate students' higher-order thinking skills (HOTS).

Design/methodology/approach: The data were collected through pre- and post-surveys as well as students' reflective journals. Further, this study employs the BERTopic model, an advanced natural language processing (NLP) technique, to conduct comprehensive topic analysis.

Findings: The results demonstrated that integrating GenAI into reading activities can positively influence students' development of HOTS. BERTopic modeling further revealed themes including intuitive immersive interaction, emotional support, assisted learning, resource access and learning convenience – reflecting students' perceptions of applying role-playing GenAI in their reading classes.

Originality/value: This study underscores the potential of role-playing GenAI to facilitate the teaching of classical literature and contributes to GenAI research in education by illustrating how teacher-tailored tools can foster HOTS within a student-centered environment. Additionally, the application of advanced topic analysis techniques provides powerful methods for understanding and exploring students' attitudes and perceptions of GenAI in language education.

ARTICLE HISTORY
Received 23 August 2025
Accepted 14 January 2026

KEYWORDS
Role-playing GenAI; HOTS development; classical literature reading; perception; BERTopic; topic modeling

Introduction

The development of higher-order thinking skills (HOTS) is widely recognized as an educational goal essential for student growth in language classes (Liu et al. 2021). HOTS involves advanced thinking abilities beyond basic knowledge, including but not limited to classification, comparison, reasoning, and abstraction (Bloom 1956; Lewis and Smith 1993). These cognitive capacities equip students to tackle complex challenges and promote lifelong learning and self-improvement (Zhan et al. 2023). Research indicates that HOTS foster academic performance, increase learning motivation, improve satisfaction, and boost classroom engagement, which offers valuable insights for curriculum design and instructional strategies (Kumpas-Lenk, Eisenschmidt, and Veispak 2018). Language and culture education,

CONTACT Sumei Wu sumeiwu@bnu.edu.cn

through effective methods such as problem-based inquiry and collaboration within active, student-centered environments (Loyens et al. 2023), fosters HOTS via dynamic knowledge construction, where students analyze and solve problems through acquiring, (re)organizing, and applying knowledge (Lu et al. 2021).

Building upon the importance of active, student-centered learning for developing HOTS, the rapid advancement of information and communications technology has made the integration of GenAI into language classes an influential approach to enhance these skills (Zhang et al. 2024). GenAI enables students to engage in Socratic-style dialogue interactions, helping them improve their reasoning abilities through questioning, critical thinking, and rebuttal (Dibek, Kursad, and Erdogan 2024). This kind of interaction is crucial for developing HOTS because it reflects the epistemological practice of how knowledge and viewpoints are constructed through argumentation, criticism, and evidence evaluation (Tang and Putra 2025). However, such interactions are difficult to achieve in traditional classrooms because it is difficult for a teacher to sustain long-term dialogue with each individual student. In addition, HOTS are often cultivated by engaging with challenging questions that lack straightforward or standard answers (Liu et al. 2021). Research highlights that GenAI presents an opportunity to quickly produce complex viewpoints for brainstorming and to provide immediate feedback on students' reasoning and argumentation. This functionality helps break students' fixed thinking patterns and stimulate creative thinking (Kim et al. 2024). As Chugh et al. (2025) articulated, when used to support, rather than replace, students' own thinking processes, GenAI can promote involvement in higher-order cognitive activities.

As an emerging educational technology, customized role-playing GenAI can simulate the personality, knowledge, language style, and behavioral patterns of specific characters, such as historical figures and literary personas. It facilitates deep, coherent, and open-ended dialogues with learners in a defined narrative framework, leveraging preset personas and maintaining continuous memory to enhance engagement and authenticity (Chen et al. 2024). Role-playing GenAI differs from other GenAI applications like intelligent tutoring systems, which primarily aims to teach specific knowledge points. Instead, it requires learners to develop their cognitive abilities (e.g. analyzing and evaluating) by engaging in the process of "experiencing and creating stories" in complex, context-rich scenarios (Kim et al. 2022, p. 18). These role-playing GenAI offer advantages in terms of initiative, consistency, and immersion, such as maintaining a persistent identity, long-term memory, and autonomous interaction, compared to the character logic input capacities of ChatGPT (Tang and Putra 2025).

While previous research has advanced our understanding of GenAI's potential in developing HOTS, much of the existing material remains theoretical and static, with limited insights into effective strategies for leveraging GenAI to cultivate HOTS in language education, particularly in a classical literature learning context. To address this gap, this study investigates the impact of a customized role-playing GenAI on learners' HOTS development and their perceptions of the learning experiences.

Literature review

The use of GenAI in language and literature education

In classical literature education, prior research has emphasized the challenges faced by the learners, including the abstract nature of literature, which hinders memory retention

and comprehension of emotional and imagery elements. Additionally, monotonous teaching methods often limit student engagement and comprehension (Wang 2025).

With technological advancements, GenAI has emerged as a transformative tool, opening new avenues for pedagogical innovation and personalized learning (Liu and Wang 2024). Specifically, GenAI supports language and classic literature instruction through customized reading guidance, simulated dialogues with virtual characters, the creation of immersive virtual scenarios using technologies such as virtual reality and the metaverse, and detailed text analysis (Baek and Hwang 2024; Chen and Wu 2024). For example, based on the video generation capability of the Runway ML Gen-3 Alpha model, Wang (2025) theoretically explored how this technology could be integrated into classical literature teaching to enhance the immersion, learning interest, and memory retention. In Chen et al.'s (2024) study, participants engaged with GenAI in a digital game, exploring the stories of the famous Chinese military strategist Sun Bin and the sage Guiguzi through role-playing. By adopting role-playing strategies, abstract classical literary works are situated in specific and meaningful contexts, enabling students to experience vivid, immersive learning. This approach promotes a deeper understanding of classical literature while encouraging analysis of character motivations, evaluation of different perspectives, and creative problem-solving, transforming interaction into active, meaningful engagement.

In recent years, researchers have increasingly focused on developing user-friendly platforms for creating GenAI applications. Tools such as ERNIE Bot, Google Dialogflow, and Poe AI are transforming the development of educational agents. For example, Kim et al. (2022) used Dialogflow to develop 'Ellie', a chatbot for language learning, demonstrating its feasibility. Similarly, Kohnke (2022) introduced a text-based chatbot for task assistance and resource recommendation, catering to diverse teaching backgrounds. These developments highlight the potential of accessible GenAI platforms to empower educators to enhance language education through innovative customized AI-driven solutions. However, current studies mainly focus on how customized AI tools can provide general learning support and feedback (Kohnke 2022; Tang and Putra 2025). The potential of GenAI as a role-playing tool has not been fully explored. Few studies have investigated how to design customized role-playing GenAI and pedagogical activities to promote HOTS in classical literature education.

Higher-order thinking skills

HOTS are essential cognitive skills that enable deeper, concept-driven understanding (e.g. Schraw et al. 2011) and involve multiple aspects such as critical thinking, problem-solving, and creativity (e.g. Araiza-Alba et al. 2021; Barta et al. 2022). Building on the importance of HOTS, Bloom's (1956) taxonomy distinguished between lower-order thinking (basic memory and knowledge) and higher-order thinking, which involves organizing, analyzing, and synthesizing information to achieve complex goals. Anderson and Krathwohl (2001) expanded upon Bloom's taxonomy by developing a revised taxonomy that categorizes cognitive skills into foundational lower-order tasks (remembering, understanding, and applying) and advanced higher-order skills (analyzing, evaluating, and creating). This revised taxonomy has become one of the most influential frameworks for understanding cognitive processes.

Empirical studies on the impact of GenAI on HOTS reveal varied outcomes. Liu and Wang (2024) demonstrated that GenAI significantly promoted critical thinking skills in English literature classes. Their findings suggested that GenAI can cultivate adaptable learning environments that support HOTS. Conversely, Essien et al. (2024) found that while GenAI improves analytical and evaluative skills, it does not significantly affect creativity. Chang et al. (2025) found that while students using ChatGPT in reflective writing courses did not show significant differences in HOTS scores compared to those in traditional learning groups, the cognitive network analysis indicated that the experimental group demonstrated more diverse conceptual understanding in their writing, as well as a better learning attitude.

Although studies have demonstrated the potential effects of GenAI on HOTS, some findings indicated that these effects did not reach statistical significance. This suggests that merely relying on technology may not be sufficient to transform the learning environment. In other words, GenAI does not automatically lead to improved student performance or the development of HOTS (Essien et al. 2024). The integration of GenAI tools in the educational environment requires careful design and research to ensure that the technology follows learning strategies and cognitive-affective principles, thereby supporting the comprehensive cognitive development of advanced skills. Additionally, the use of customized role-playing GenAI to cultivate HOTS remains underexplored, especially in the classical literature education setting. To address the gap, this study investigates the impact of a customized role-playing GenAI, developed via AgentBuilder, on learners' cognitive skills. The research questions include:

(1) In what ways does customized role-playing GenAI influence the development of HOTS in classical literature education?
(2) What are the students' attitudes and perceptions toward the use of customized role-playing GenAI?

Method

Participants

The educational intervention was conducted over a five-week middle school classical literature reading course. The program involved 20 students (55% male and 45% female) from Chongqing, China, all of whom were secondary pupils aged 14–15. All the participants had already received some initial exposure to classical novels during their primary and early middle school years, and possessed the general language skills required to complete the courses. This innovative approach aimed to deepen students' understanding of classic literature and support language learning, while fostering the development of students' HOTS.

Procedure

This study adopted a mixed-methods approach to address the research questions. Students first completed a HOTS pre-survey. They then participated in a five-week classical literature reading project, allocating 40 min per week, which incorporated role-playing with GenAI. The teaching materials were carefully selected from four Chinese classical

masterpieces, including *Lu Zhishen Raises Havoc in Taohua Village, Lin Daiyu Enters the Jia Mansion, Three Visits to the Thatched Cottage* and *Sun Wukong and the First Attempt to Borrow the Banana Fan*. At the end of the program, students completed a post-survey to evaluate changes in their HOTS and submitted reflective journals to document their learning experiences and perceptions.

Drawing on major previous research focusing on HOTS cultivation (Anderson and Krathwohl 2001; Bloom 1956; Chang et al. 2025), we designed HOTS training activities, which were structured into four phases: thinking, discussion, interaction, and application, as presented in Figure 1 and Table 1. In the thinking stage, students independently engaged with the texts, deconstructing them through reading and responding to a series of questions based on Bloom's taxonomy. This stage aimed to enhance their analysis and evaluation abilities. The second phase involved small group discussions with three students each. Through argumentation and reaching consensus on HOT-related questions, learners collaboratively constructed knowledge and sharpened their critical thinking skills. In the third stage, students experienced immersive learning facilitated by role-playing with GenAI, receiving instant feedback. They were prompted to formulate questions to interact with GenAI and evaluate its outputs, fostering deeper engagement and evaluative thinking. They raised questions to Lin Daiyu's AI agent, such as, 'When you first entered the Jia household, you watched others before drinking tea. If someone deliberately made a mistake to tease you, would you imitate them?' 'If you came to the modern world, what would you want to do?'. The final phase focused on application. Students completed HOTS tasks that required them to reflect on their understanding of the texts. These tasks included analyzing role behaviors, evaluating the ethical values of events, and imagining alternative plot developments. This process aimed to cultivate their problem-solving and creative thinking abilities.

The process of building role-playing GenAI

In this study, we designed GenAI tools to engage students with classical literature. We evaluated multiple platforms and ultimately chose AgentBuilder due to its user-friendly interface and accessibility. By adopting a prompt-based approach, teachers can

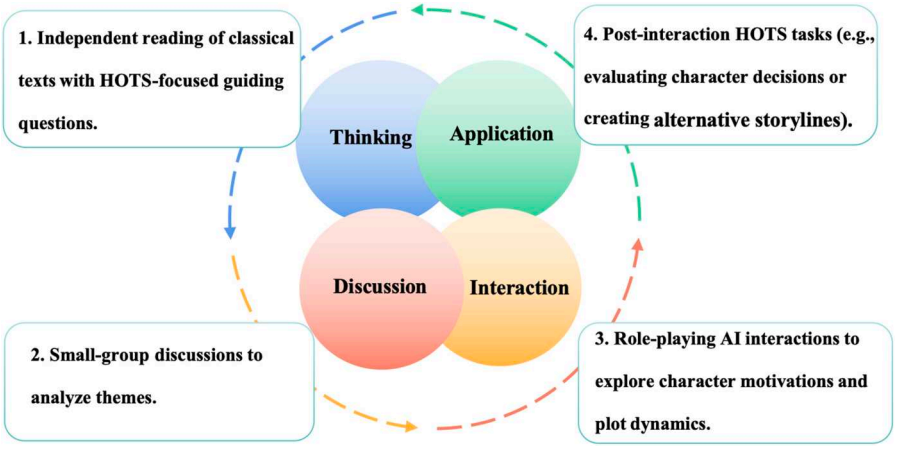

Figure 1. Learning cycle.

Table 1. The four phases of HOTS cultivation activities.

Stages	Stage tasks and objective	Guiding questions for HOTS development
Thinking	Learners independently read texts and reflect on the given HOTS training questions. This stage aims to enhance their analytical and evaluation capabilities.	Wang Xifeng enters the scene 'boisterously and without decorum', while Daiyu acts 'cautiously in every step'. How do these contrasting behaviors reflect their positions and personalities within the Jia family?
Discussion	Learners engage in group discussions on these questions, collaboratively constructing knowledge and enhancing collaboration, communication, and critical thinking skills. This stage aims to enhance their analytical and evaluation capabilities.	Daiyu tells Grandmother Jia she 'only just began reading the Four Books', but later tells Baoyu she 'never studied'. Why does she hide her education? Is this a wise survival strategy or a surrender to societal expectations?
Interaction	Students are encouraged to think critically and pose questions during the role-playing interactions with GenAI, then evaluate outputs, which promotes student engagement and reflection on their learning. This stage aims to enhance their analytical, evaluation and creative capabilities	Suppose the scabby-headed monk's prophecy came true (Daiyu never meets the Jia family). Craft an alternative destiny for her.
Application	Students complete HOTS tasks that require reflection on text understanding, involving analyzing character actions, evaluating ethical values of events, and imagining alternative plot developments. Learners are required to solve questions and be creative. This stage aims to enhance their analytical, evaluation and creative capabilities	

effortlessly develop AI-driven role-playing scenarios with minimal coding, significantly lowering the development barrier. It also excels in maintaining role consistency during long conversations without deviating from the identity and knowledge background.

For this project, the GenAI was endowed with a role-playing function, allowing it to assume specific characters and interact vividly with students. These characters were not fictional inventions but closely aligned with the literary figures from the texts studied in the course. For example, in *Dream of Red Mansions*, one of China's four great classical novels, we created a GenAI tool representing the main character, Lin Daiyu. As shown in Figure 2, the process began by selecting a language model, then defining the avatar's name/role (e.g. Lin Daiyu) and creating an introduction to develop an interaction context. Next, persona logic was developed to simulate Lin Daiyu's dialogue style, using tools like image-generation plugins for scene visuals and voice options to match the character. Finally, parameters such as auto-follow-up questions, internet search functions, and chat history were configured to shape the GenAI's role-play for literary exploration. In other words, the customized process involved setting up roles and scenarios for role-playing according to classical literature to tailor interactions for the enhancement of literary understanding and engagement.

To address ethical concerns, all data is stored in an anonymized and encrypted manner. We also reviewed the initial dialogues of the AI characters multiple times and conducted multiple rounds of iterative optimization based on the advice of the educator and expert to ensure the AI roles' cultural expression accuracy.

Instrument

HOTS questionnaire

This study adopted the questionnaire developed by Hwang et al. (2017) to measure students' HOTS development. It consisted of 11 items, covering three dimensions: problem-

Figure 2. The process of building role-playing GenAI.

solving (4 items), critical thinking (4 items), and creativity (3 items). The questionnaire used a five-point Likert scale (1 = 'strongly disagree', 5 = 'strongly agree') to gauge the different degrees of learners' HOTS.

To ensure the quality and appropriateness of this scale for Chinese students, the content validity of the questionnaire was evaluated and reviewed by two experts in language and educational technology, and piloted with two Chinese middle school students. They agreed that the content validity of the questionnaire was generally good and provided suggestions to improve the comprehensibility. The reliability of the questionnaire in the present study was also examined, which showed that it was acceptable (Cronbach's $\alpha = 0.748$).

The instrument was chosen for its comprehensive coverage of essential HOTS and its suitability for capturing changes in students' higher-order cognitive abilities within the context of innovative learning activities.

Student reflective journal

Regarding the collection of qualitative data, students were asked to reflect on their perceptions and experiences of learning through role-playing with GenAI. The reflective journal in the study was primary based on the SWOT model (Chen et al. 2024), including four aspects, including (1) the benefits gained from using role-playing GenAI, (2) the limitations or shortcomings of this tool, (3) the opportunities that this tool offered for classical literature learning and HOTS cultivation, and (4) the challenges encountered. Additionally, in order to provide students with more insights for writing the reflective journals, the researchers also set prompts like the differences between integrating role-playing GenAI into classical literature learning and HOTS cultivation and other learning methods, and their personal feelings and overall impression of using role-playing GenAI.

Data analysis

As the assumptions of normality were satisfied using the Shapiro-Wilk test ($p > 0.05$), the paired-sample t-test was conducted to compare students' pre- and post-intervention HOTS, aiming to determine whether the integration of role-playing GenAI activities in classical literature learning could improve HOTS. The analysis was performed using SPSS version 26.0.

For the qualitative analysis of students' reflective journals, the BERTopic model – a sophisticated NLP technique – was employed to identify key themes. This approach was selected because, as a novel modern topic modeling method, it can extract hidden thematic patterns and features in the text, playing a crucial role in mining massive complex text data. Compared to traditional topic modeling methods, such as Latent Dirichlet Allocation (LDA), which typically treat documents as collections of words, easily ignoring subtle nuances of word meanings and sentence context, the BERTopic model, as a text embedding approach, effectively addresses these deficiencies (Egger and Yu 2022). It enables an in-depth exploration of participants' attitudes, perceptions, and insights regarding their experiences with the GenAI-enhanced HOTS reading program. The process began with preprocessing the text to remove irrelevant elements, such as punctuation, special characters, and stop words. Next, document embeddings were generated using the all-MiniLM-L6-v2 model. Dimensionality reduction was then

applied with UMAP to enhance the accuracy of topic detection. Finally, BERTopic utilized HDBSCAN for clustering, followed by extracting topics through the c-TF-IDF algorithm.

To ensure theme reliability, the research team conducted a manual review after BERTopic generated topics. Two researchers independently examined the high-frequency words, representative statements, and themes, and then negotiated to refine the themes for better accuracy in reflecting the students' reflective journal contents.

Results

RQ 1: In what ways does customized role-playing GenAI influence the development of HOTS in classical literature education?

A paired T-test was conducted to examine the impact of the teaching intervention on students' HOTS (see Table 2). The results showed a significant improvement in HOTS scores ($p < 0.01$). Specifically, the mean score of the post-test (M = 3.49, $SD = 0.32$) was notably higher than that of the pre-test (M = 3.32, $SD = 0.30$), with $t = -3.483$ and $p = 0.002$. These findings suggest that the role-playing GenAI-based teaching method effectively promoted the development of students' HOTS.

To confirm the finding, we also conducted the Wilcoxon signed-rank test. The results revealed that role-playing GenAI-based teaching intervention led to a statistically significant improvement in students' HOTS scores ($Z = -2.772$, $p = 0.006$).

RQ 2: What are the students' attitudes and perceptions toward the use of customized role-playing GenAI?

Using BERTopic, four main themes and one outlier theme emerged from the analyzed textual data, collectively depicting the varied experiences of language learners engaging with the role-play GenAI, as shown in Figure 3. The figure presents the c-TF-IDF scores of keywords within each theme, highlighting the importance of these terms in representing their respective themes.

By conducting a detailed analysis of the top ten keywords associated with each topic and examining the original textual data (see Table 3), we have identified and labeled the themes accordingly. The detailed interpretations of these themes are as follows.

Topic – 1: exhibiting intelligent efficiency and operational complexity

This theme focused on users' experiences with the operational aspects and potential challenges of using role-playing GenAI. Keywords such as 'intelligent', 'operation' and 'beneficial' indicated that learners generally perceived GenAI as intelligent and convenient, recognizing its role in enhancing learning efficiency and optimizing knowledge

Table 2. Comparing students' HOTS before and after the intervention.

	Mean	*SD*	Paired T-test *t*-value (*p*)	Wilcoxon test *Z*-value (*p*)
Pre-survey	3.32	0.30	−3.483(0.002**)	−2.772(0.007**)
Post-survey	3.49	0.32		

*$p < 0.05$ **$p < 0.01$.

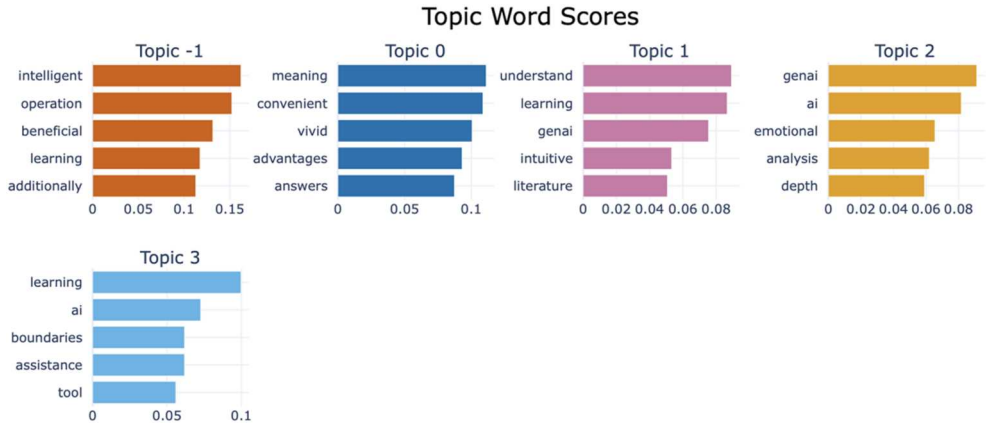

Figure 3. Topic words generated by BERTopic.

acquisition. However, feedback within this theme also highlighted limitations, including the complexity of the operation process, concerns about the completeness and accuracy of GenAI responses, and a lack of emotional depth. Although the sample size for this theme was small, its findings provided valuable guidance for improving the design and usability of educational GenAI applications.

Topic 0: enhancing learning convenience and enriching reading experience

This theme highlighted learners' positive perceptions of GenAI's role in enhancing learning convenience, enriching literary reading experiences, and making interactions more engaging. For instance, by transforming abstract texts into vivid scenes, role-playing GenAI was reshaping learner-text relationships, turning passive readers into active co-creators of meaning. As S20 expressed, 'When I talk to Dai Yu [the literary figure], I feel that I am both a reader and an accomplice of Xueqin Cao [the writer]'. Some feedback

Table 3. Keywords for each topic identified by BERTopic.

Topic	Topic name	Topic keywords	Representative docs
−1	Exhibiting intelligent efficiency and operational complexity	['intelligent', 'operation', 'beneficial', 'learning', 'additionally', 'efficiency', 'inaccuracies', 'improved', 'incomplete', 'frustrating']	GenAI has improved my learning efficiency and has been beneficial. It is intelligent and convenient, yet the operation can be troublesome …
0	Improving the learning convenience and reading experience richness	['meaning', 'convenient', 'vivid', 'advantages', 'answers', 'classical', 'literature', 'genai', 'questions', 'experience']	GenAI creates meaning and actively participating, and transforms passive readers into active creators of meaning …
1	Fostering literature understanding and intuitive interaction	['understand', 'learning', 'genai', 'intuitive', 'literature', 'questions', 'challenges', 'classical', 'beneficial', 'understanding']	The role-playing feature of GenAI enables me to have an intuitive, vivid and immersive experience …
2	Enhancing emotional support and role-playing effects	['genai', 'ai', 'emotional', 'analysis', 'depth', 'playing', 'role', 'understanding', 'learning', 'similarity']	It was like meeting the characters face to face … A drawback is its lack of genuine emotional expression …
3	Promoting knowledge expansion and resource acquisition	['learning', 'ai', 'boundaries', 'assistance', 'tool', 'knowledge', 'fast', 'concise', 'process', 'text']	GenAI provides resources not found in traditional textbooks … It breaks the boundaries of traditional learning …

noted minor delays in GenAI response times; however, overall, users rated the experience as 'greatly enhancing the efficiency and interest of participating in the learning process'. These tools helped students understand abstract concepts through text analysis and visual representations of classical literature. For instance, S19 commented, 'Talking with GenAI helps me better understand the content, and when I encounter words that I don not understand, it can also help me analyze'.

Topic 1: fostering literature comprehension and facilitating intuitive interaction

This theme centered on learners' perceptions of GenAI's role in facilitating understanding, deepening engagement, and making the learning process more intuitive. Many users believed that GenAI provided rapid, detailed responses, particularly highlighting its strengths in literary reading and text analysis. For example, S9 remarked that 'in-depth analysis of questions they don't understand' helped deepen their material understanding. Further, GenAI showed promise in stimulating interest and increasing classroom participation; students found that role-playing activities made classical studies more lively. However, some users experienced interface complexity and occasional misunderstandings during initial interactions, suggesting a need for developers and educators to optimize user guidance and streamline operational procedures to improve the learning experience.

Topic 2: enhancing emotional support and role-playing experience

This theme concentrated on learners' feedback concerning GenAI's emotional support and role-playing capabilities. Students generally found that GenAI simulated real interactions and demonstrated patience – even when faced with repetitive questions – alleviating stress from traditional methods. Role-playing with GenAI allowed users to 'embody' the affective states of characters, generated emotional connections, and enhanced interactive learning. For instance, S5 mentioned feeling as if he could genuinely feel the emotions of different characters, while S2 found the experience akin to meeting characters face to face. However, some feedback indicated that GenAI currently lacked authentic emotional expression and subtle emotional communication that hinder deeper immersion. This suggested a learner expectation for more 'humanized' interactions, indicating that future improvements should enhance GenAI's emotional expressiveness to create more natural and relatable interactions.

Topic 3: promoting knowledge expansion and resource acquisition

This theme highlighted GenAI's role in supporting learning, expanding knowledge horizons, and transcending traditional educational boundaries. For most students, role-playing with GenAI has been a positive experience, serving as a valuable complement to traditional textbooks and enriching the learning process. Many reported that GenAI expanded resource access, especially in literature studies, where it could quickly provide the relevant information. The efficiency of role-play GenAI was widely recognized: for example, S3 reflected in his reflective journal that using GenAI was easy and practical. The interface was generally viewed as efficient and straightforward, facilitating problem-solving and interactive engagement. However, some feedback pointed to areas for improvement, such as limited flexibility in operation and the need for a more intuitive interface. These insights implied that stakeholders should focus on enhancing

personalization and user experience without compromising operational efficiency, making GenAI a more adaptable and user-friendly educational tool.

In addition to identifying specific topics, BERTopic generates various graphical representations to enhance interpretability. The graphical similarity matrix displays topic similarities as heatmaps. As displayed in Figure 4, Topic 2 and Topic 3 had a high similarity, and Topic 1 was also connected to Topic 2, with a similarity score of 0.86. This indicated that the emotional support provided by GenAI and the role-playing learning experience were linked to the cognitive aspect of learning. Findings suggest that learners value emotional interaction and role-playing as complementary tools for knowledge acquisition in literary studies, emphasizing the need for emotional and innovative design elements to boost cognitive engagement and learning efficiency.

Hierarchical clustering shows the relationships among the identified topics. As illustrated in Figure 5, if aiming to consolidate themes, there was a tendency to merge Topic 1 and Topic 3, as both reflect GenAI's role as a cognitive assistive tool, highlighted by overlapping keywords such as 'learning', 'understanding', and 'assistance'. In contrast, Topic 2 emphasized emotional support and human-computer interaction, while Topic 1 pertains to technical operations related to interfaces and processes. This clustering underscored the distinct yet interconnected aspects of GenAI's role in educational contexts.

The Intertopic Distance Map (see Figure 6), which demonstrated the degree of proximity or distance between the topics, displayed themes evenly distributed within a two-dimensional space without overlap, indicating that these dimensions are distinct and non-substitutable. The map further demonstrated that GenAI educational applications

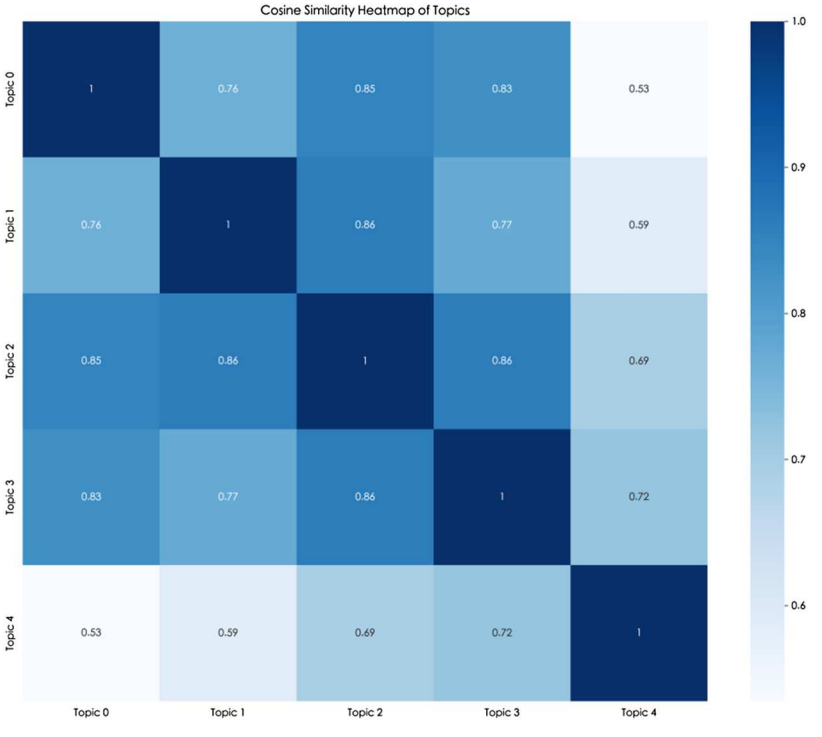

Figure 4. Similarity matrix visualization.

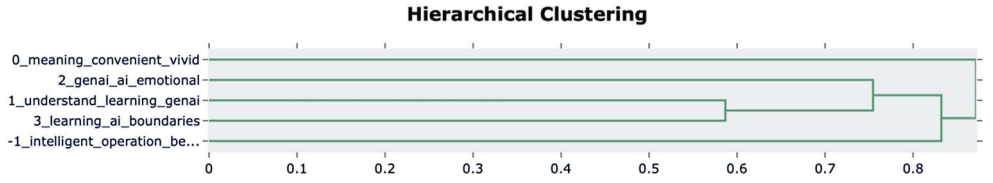

Figure 5. Hierarchical clustering.

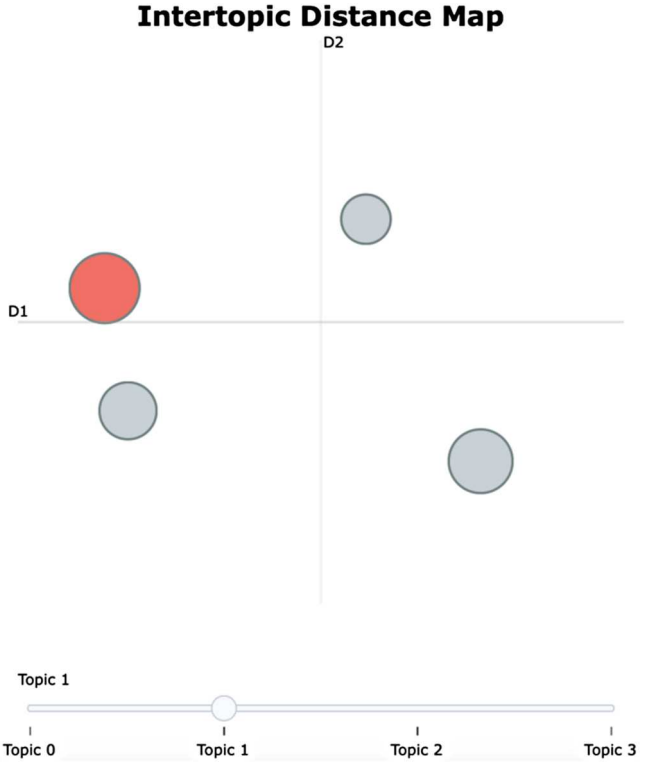

Figure 6. Intertopic distance map visualization.

serve multiple functions: facilitating learning and resource acquisition (Topic 3), supporting intuitive interaction and cognitive construction (Topic 1), enhancing convenience and overall learning experience (Topic 0), and providing emotional support (Topic 2). This multidimensionality underscores the comprehensive role of GenAI in enriching the educational experience.

Discussion

Classical literature learning is vital in language classrooms. Despite teachers' efforts, the traditional instructor-centered environment is often confined to written texts and direct reading practice that fall short of active engagement and interaction. With minimal opportunities to fully experience the literary texts, traditional classical literature learning

poses a significant challenge for students to understand the cultural context, maintain learning interest and comprehend complex narratives and linguistic structures (Baek and Hwang 2024). In this situation, the integration of GenAI into classical literature instruction presents a promising avenue. Given the indispensable role of HOTS in GenAI use and language education, it is imperative that GenAI set its sights on meaningful educational purposes and cultivate students' higher cognition to support profound and lifelong learning and growth, instead of being limited to solely focusing on improving scores in standardized academic assessments. The present study contributes intriguing findings and insights into the impact of the interaction of role-playing GenAI on HOTS. Results highlight the effectiveness of role-playing GenAI in creating an adaptive and interactive language learning environment tailored to HOTS development, corroborating Fabio, Plebe, and Suriano (2024), who advocate that the conscious and informed adoption of GenAI has introduced innovative possibilities for facilitating HOTS and learning processes, due to its capabilities to promote critical analysis and reflection through intricate scenarios and questions.

Analytical skills, as an important component of HOTS, mainly benefit from the role-playing GenAI being a powerful cognitive tool. Unlike generic AI tools, the customized design of role-playing GenAI in this study provides a unique scaffolding approach for literature learning within language classes, particularly through the simulation of classical characters. This aligns with Chen et al. (2024), who emphasized that integrating GenAI into EFL role-playing activities can enhance engagement, making generated content easier for students to comprehend. Our research suggests a key pedagogical implication of findings is that role-playing GenAI interactions allow students to reinterpret classical narratives and construct an innovative, student-centered learning pathway for cultural inheritance (Loyens et al. 2023). This innovative teaching practice not only expands learning resources but also integrates character personalities, plotlines, and the status and power of roles, guiding students to move beyond the text and conduct causal logic analysis, which provides a more exploratory analytical path.

Regarding evaluation, by simulating literary characters, who have stances and values in the context of the work and historical background, role-playing GenAI tools generate cognitive empathy and conflict through interaction with students. The program allows learners to build textual knowledge through immersive experiences, reducing emotional distance from classic texts. By integrating visual and other sensory cues, abstract literary symbols are transformed into immediate figurative, and interactive experiences (Chen and Wu 2024). Moreover, students are also required to consider multiple factors, such as historical context, family situation and personal personality, when making a critical evaluation. In contrast to conventional teacher-fronted environments, where students often find it difficult to form their independent value judgments for the dominant role of the teacher's interpretation and textbook conclusion, this pedagogy empowers students to break through single authoritative interpretations and explore multiple perspectives. In this study, as an indefatigable debate partner, GenAI becomes an important tool for promoting critical thinking and evaluation skills among students.

Role-playing GenAI also responds to students' creative prompts and provides contextualized feedback in a way that conforms to the character's setting for learners' creativity cultivation. Traditional classrooms often lack sufficient opportunities for innovative expressive activities and creative practice, often confined to imitating and reciting

famous quotes from memory (Baek and Hwang 2024). Reflection journals show that students transition from passive readers to active participants, exhibiting the potential to overcome the cognitive limitations of focusing on a single text. This shift introduces new possibilities for creative expression in language classes.

When applying this pedagogy, we also illustrate limitations in conveying emotional authenticity, which may lead to superficial engagement, partly due to current shortcomings of large language models (LLMs), lacking the nuanced emotional capabilities. As discussed, simulated emotional expressions often fail to replicate human subtleties, leading to superficial or inaccurate portrayals within role-playing environments (Chen et al. 2024). Further exploration of how LLMs could be improved – such as incorporating affective computing – would be valuable. Moreover, participants' reflective data suggested issues such as over-reliance on GenAI, which raises concerns about cognitive shortcuts that might discourage students from critically evaluating sources. To address these challenges, stakeholders should clarify the role of GenAI in education and explore innovative teaching models that combine GenAI tools with teacher guidance and peer collaboration.

Conclusion and implications

This study investigated the impact of customized role-playing GenAI on learners' HOTS within a classical literature reading program, while also exploring their perceptions and experiences. Findings indicated that integrating role-playing GenAI positively affected learners' HOTS. BERTopic topic modeling highlighted GenAI's advantages in efficiency, convenience, and immersive engagement, and challenges related to emotional connection and operational complexity. These findings offer valuable empirical insights to guide future educational applications, emphasizing the need to optimize GenAI's emotional expressiveness and usability to maximize its pedagogical potential.

The present study offers significant theoretical implications for the integration of GenAI in education. It reveals how students co-construct knowledge through interactions with role-playing GenAI, fostering skills in analysis, evaluation, and creativity. It also reveals that the digital embodiment of characters can enhance cognitive engagement and empathetic understanding. This research positions customized GenAI not merely as a tool but as an active participant and collaborator in the cognitive and social processes of learning, providing a new theoretical perspective for understanding human-computer collaboration in language and literature education contexts.

This study holds critical implications for educational practice. From a practical standpoint, this program offers empirical evidence of HOTS cultivation within GenAI-assisted literature reading classes, presenting an innovative instructional design case for educators to achieve pedagogical goals and cultivating HOTS in language education. Specifically, the findings affirm the potential of customized role-playing GenAI to enhance HOTS by enabling GenAI to create meaning and participate in the learning process – transforming passive readers into active meaning-makers. By simulating literary characters, the integration of GenAI in reading activities helps bridge the gap between historical narratives and modern learners. In addition, this study provides a pedagogical example of user-friendly, customized GenAI, empowering instructors with limited programming experience to develop context-specific GenAI tools for engaging and effective teaching design. To translate these findings into more valuable pedagogical practice, teacher

scaffolding could include demonstrating how to critically evaluate GenAI-generated texts and highlighting limitations, such as inaccuracies, while progressively guiding students to identify issues independently. Additionally, facilitating student discussions about their analyses with GenAI can promote critical dialogue, encouraging learners to articulate reasoning processes rather than passively accepting GenAI outputs. Schools might also implement training on effective GenAI usage, teaching students to craft precise prompts for more analytical responses.

The limitations of this study should be acknowledged. One notable limitation is the lack of data triangulation. Future research could incorporate broader and diverse data collection approaches to provide a more comprehensive understanding. Additionally, the five-week duration restricts capturing the cumulative development of HOTS, which requires sustained practice, and may overlook shifts in student perceptions or engagement. Longitudinal studies are recommended to track long-term skill development and associated effects. Furthermore, given the limitations of the participant sample size, the quantitative data in this study primarily serve as preliminary effect validation. Future studies would benefit from recruiting more diverse samples across multiple institutions or educational contexts to enhance the robustness and applicability of the results. Despite these constraints, this educational innovation contributes to practical teaching methods amid rapid technological advancements. It offers insights into how GenAI can be leveraged to cultivate HOTS in language education and inject new vitality into classical literature reading, thereby opening promising avenues for pedagogical innovation and exploration.

Acknowledgements

This study was supported by a grant from the Fundamental Research Funds for the Central Universities in China (310400209512). We would like to extend our great appreciation for the anonymous reviewers and editors for their constructive feedback on the manuscript.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Zhaoyang Xiong http://orcid.org/0009-0006-6934-9675
Junjie Gavin Wu https://orcid.org/0000-0003-4937-4401
Sumei Wu https://orcid.org/0000-0001-7626-5927
Junhua Xian https://orcid.org/0009-0001-0585-7535

References

Anderson, L. W., and D. R. Krathwohl. 2001. *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives: Complete Edition*. New York: Longman.

Araiza-Alba, P., T. Keane, W. S. Chen, and J. Kaufman. 2021. "Immersive Virtual Reality as a Tool to Learn Problem-solving Skills." *Computers & Education* 164:104121. https://doi.org/10.1016/j.compedu.2020.104121.

Baek, J., and Y. Hwang. 2024. "Old Tales, New Tools: How AI Chatbots and the Metaverse Are Transforming Premodern Classical Literature Reading Activities." *International Journal of Educational Research* 128:102461. https://doi.org/10.1016/j.ijer.2024.102461.

Barta, A., L. A. Fodor, B. Tamas, and I. Szamoskozi. 2022. "The Development of Students Critical Thinking Abilities and Dispositions through the Concept Mapping Learning Method – a Meta-analysis." *Educational Research Review* 37:100481. https://doi.org/10.1016/j.edurev.2022.100481.

Bloom, B. S. 1956. *Taxonomy of Educational Objectives: The Classification of Educational Goals, Handbook 1: Cognitive domain*. New York: David McKay.

Chang, C. Y., H. C. Lin, C. Yin, and K. H. Yang. 2025. "Generative AI-assisted Reflective Writing for Improving Students' Higher Order Thinking: Evidence from Quantitative and Epistemic Network Analysis." *Educational Technology and Society* 28 (1): 270–285. https://doi.org/10.30191/ETS.202501_28(1).TP03.

Chen, A., J. Jia, Y. Li, and L. Fu. 2024. "Investigating the Effect of Role-play Activity with GENAI Agent on EFL Students' Speaking Performance." *Journal of Educational Computing Research* 63 (1): 99–125. https://doi.org/10.1177/07356331241299058.

Chen, X., and D. Wu. 2024. "Automatic Generation of Multimedia Teaching Materials Based on Generative AI: Taking Tang Poetry as an Example." *IEEE Transactions on Learning Technologies* 17:1327–1340. https://doi.org/10.1109/tlt.2024.3378279.

Chen, Z., Xiong, Z., Ruan, X., Jiang, S., Wei, W., & Fang, K. 2024. "Exploring Learners' Interactions with GenAI Agents in Educational Games: Typologies and Emotional Factors in Human-Computer Interaction." In *2024 4th International Conference on Educational Technology (ICET)*, 136–140. IEEE. https://doi.org/10.1109/ICET62460.2024.10868064.

Chugh, R., D. Turnbull, S. Kutty, F. Sabrina, M. M. Rashid, A. Morshed, S. Azad, S. Kaisar, and S. Subramani. 2025. "Generative AI as a Learning Assistant in ICT Education: Student Perspectives and Educational Implications." *Education and Information Technologies* 30 (16): 23693–23728. https://doi.org/10.1007/s10639-025-13686-3.

Dibek, M. I., M. S. Kursad, and T. Erdogan. 2024. "Influence of Artificial Intelligence Tools on Higher Order Thinking Skills: A Meta-analysis." *Interactive Learning Environments* 33 (3): 2216–2238. https://doi.org/10.1080/10494820.2024.2402028.

Egger, R., and J. Yu. 2022. "A Topic Modeling Comparison between LDA, NMF, Top2VEC, and BERTopic to Demystify Twitter Posts." *Frontiers in Sociology* 7:886498. https://doi.org/10.3389/fsoc.2022.886498.

Essien, A., O. T. Bukoye, X. O'Dea, and M. Kremantzis. 2024. "The Influence of AI Text Generators on Critical Thinking Skills in UK Business Schools." *Studies in Higher Education* 49 (5): 865–882. https://doi.org/10.1080/03075079.2024.2316881.

Fabio, R. A., A. Plebe, and R. Suriano. 2024. "AI-based Chatbot Interactions and Critical Thinking Skills: An Exploratory Study." *Current Psychology* 44: 8082–8095. https://doi.org/10.1007/s12144-024-06795-8.

Hwang, G., C. Lai, J. Liang, H. Chu, and C. Tsai. 2017. "A Long-term Experiment to Investigate the Relationships between High School Students' Perceptions of Mobile Learning and Peer Interaction and HOTS Tendencies." *Educational Technology Research and Development* 66 (1): 75–93. https://doi.org/10.1007/s11423-017-9540-3.

Kim, H., H. Yang, D. Shin, and J. H. Lee. 2022. "Design Principles and Architecture of a Second Language Learning Chatbot." *Language Learning & Technology* 26 (1): 1–18. https://doi.org/10.64152/10125/73463.

Kim, J., S. Yu, R. Detrick, and N. Li. 2024. "Exploring Students' Perspectives on Generative AI-assisted Academic Writing." *Education and Information Technologies* 30 (1): 1265–1300. https://doi.org/10.1007/s10639-024-12878-7.

Kohnke, L. 2022. "A Pedagogical Chatbot: A Supplemental Language Learning Tool." *RELC Journal* 54 (3): 828–838. https://doi.org/10.1177/00336882211067054.

Kumpas-Lenk, K., E. Eisenschmidt, and A. Veispak. 2018. "Does the Design of Learning Outcomes Matter from Students' Perspective?" *Studies in Educational Evaluation* 59:179–186. https://doi.org/10.1016/j.stueduc.2018.07.008.

Lewis, A., and D. Smith. 1993. "Defining Higher Order Thinking." *Theory into Practice* 32 (3): 131–137. https://doi.org/10.1080/00405849309543588.

Liu, J., Y. Ma, X. Sun, Z. Zhu, and Y. Xu. 2021. "A Systematic Review of HOTS by Visualizing Its Structure through HistCite and CiteSpace Software." *The Asia-Pacific Education Researcher* 31 (6): 635–645. https://doi.org/10.1007/s40299-021-00614-5.

Liu, W., and Y. Wang. 2024. "The Effects of Using AI Tools on Critical Thinking in English Literature Classes among EFL Learners: An Intervention Study." *European Journal of Education* 59 (4). e12804. https://doi.org/10.1111/ejed.12804.

Loyens, S. M. M., J. E. Van Meerten, L. Schaap, and L. Wijnia. 2023. "Situating Higher-order, Critical, and Critical-analytic Thinking in Problem- and Project-based Learning Environments: A Systematic Review." *Educational Psychology Review* 35 (2): 39. https://doi.org/10.1007/s10648-023-09757-x.

Lu, K., H. H. Yang, Y. Shi, and X. Wang. 2021. "Examining the Key Influencing Factors on College Students' HOTS in the Smart Classroom Environment." *International Journal of Educational Technology in Higher Education* 18 (1): 1. https://doi.org/10.1186/s41239-020-00238-7.

Schraw, G., M. T. McCrudden, S. Lehman, and B. Hoffman. 2011. "An Overview of Thinking Skills." In *Assessment of Higher Order Thinking Skills*, edited by G. Schraw and D. R. Robinson, 19–45. Charlotte: Information Age Publishing.

Tang, K., and G. B. S. Putra. 2025. "Generative AI as a Dialogic Partner: Enhancing Multiple Perspectives, Reasoning, and Argumentation in Science Education with Customized Chatbots." *Journal of Science Education and Technology*. https://doi.org/10.1007/s10956-025-10240-1.

Wang, Y. 2025. "Prospects of Generative Video Technology in the Field of Chinese Classical Literature: Exploring Female Education through the Reimagining of Chinese Poetic Imagery Using Runway ML Gen-3 Alpha." In *2025 5th International Conference on Artificial Intelligence and Education (ICAIE)*, 407–411. https://doi.org/10.1109/icaie64856.2025.11158710.

Zhan, Y., Z. Yan, Z. H. Wan, X. Wang, Y. Zeng, M. Yang, and L. Yang. 2023. "Effects of Online Peer Assessment on Higher-order Thinking: A Meta-analysis." *British Journal of Educational Technology* 54 (4): 817–835. https://doi.org/10.1111/bjet.13310.

Zhang, D., J. G. Wu, and Z. Fu. 2024. "From Shy to Fly: Facilitating EFL Learners' Willingness to Communicate with an AI Chatbot and an Intelligent Tutoring System." *System* 127: 103501. https://doi.org/10.1016/j.system.2024.103501.