

Enhancing nursing education: An AI-powered Chatbots for fostering engagement and higher-order thinking skills

Ruowei Wang

Shandong Xiehe University

Arumugam Raman

arumugam@uum.edu.my

Universiti Utara Malaysia

Research Article

Keywords: AI-powered chatbots, Student engagement, Higher-order thinking skills (HOTS), Perceived usefulness (PUC), Ease of use (EoU), Feedback quality (FBQ), Self-regulated learning (SRL)

Posted Date: May 8th, 2025

DOI: <https://doi.org/10.21203/rs.3.rs-6372448/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

AI-powered chatbots are increasingly integrated into educational settings, offering opportunities to enhance student engagement and higher-order thinking skills (HOTS). However, limited research exists on their role in nursing education, especially in China. This study aimed to explore how AI-powered chatbots impact nursing students' engagement and the development of HOTS, mediated by feedback quality (FBQ) and self-regulated learning (SRL).

Methods

A cross-sectional, quantitative research design was employed to investigate the interplay between perceived usefulness (PUC), ease of use (EoU), engagement (ENG), FBQ, SRL, and HOTS. 470 nursing students from different academic years participated in the study. Data were collected using a structured survey measuring six key constructs. Partial Least Squares Structural Equation Modelling (PLS-SEM) was employed to evaluate the direct and indirect relationships within the conceptual framework.

Results

PUC and EoU significantly influenced ENG, which, in turn, strongly mediated the effects of FBQ and SRL on HOTS. ENG had a substantial impact on FBQ ($\beta = 0.852$) and SRL ($\beta = 0.892$), while the combined indirect effect of FBQ and SRL on HOTS ($\beta = 0.798$) demonstrated the critical role of these mediators. The study also confirmed that intuitive chatbot design and high-quality, timely feedback are essential for fostering cognitive skills.

Conclusions

AI-powered chatbots show promise in enhancing engagement and supporting the development of higher-order thinking skills in nursing education. The findings emphasize the need for scalable, user-friendly chatbot systems tailored to educational contexts. Future research should focus on advanced feedback algorithms, long-term impacts on clinical competency, and scalability in diverse learning environments.

Background

Chatbots in education have gained significant momentum in recent years, thanks to advancements in artificial intelligence (AI) that have made these tools more interactive and adaptable for teaching and learning [1]. In healthcare education—where knowledge is vast and constantly evolving—chatbots offer a unique solution by providing students with on-demand access to resources, information, and

personalized support [2–4]. Nursing education, in particular, benefits from these innovations, as nursing students must balance mastering theoretical concepts with acquiring practical skills to prepare for their demanding roles.

In China, the demand for skilled healthcare professionals has surged due to ageing populations, rapid urbanization, and a growing burden of chronic diseases [5, 6]. As a result, nursing education is under pressure to produce highly competent graduates who can meet these challenges. While traditional teaching methods have been effective, they often require students to quickly learn a large amount of complex knowledge and passively absorb knowledge points, and may fall short in addressing students' personalized learning needs and the cultivation of higher-order thinking skills [7, 8]. With their ability to deliver instant, individualized feedback and support, Chatbots present a promising supplementary or auxiliary solution [9–11].

In addition, nursing students are faced with the continuous updating of medical and health knowledge. In this context, chatbots have proven effective in fostering self-directed learning, enhancing engagement, and promoting critical thinking skills [12–16]. In China, the adoption of AI chatbots in nursing education is gradually growing as educational institutions leverage technology to improve student outcomes [17].

AI-powered virtual assistants are increasingly being integrated into healthcare education to enhance student learning and engagement. For example, iFlytek, a leading AI company, has developed advanced educational tools such as the T10 AI Learning Machine, which provides personalized learning experiences and AI-assisted grading in subjects like English and medical sciences [18]. While Xiaoyi, an AI developed by iFlytek, made history by passing China's National Medical Licensing Examination in 2017, its primary role is to assist doctors in analyzing patient information rather than directly training nursing students [19]. Additionally, AI-driven platforms are being used to support nursing education through interactive tutoring systems, personalized quizzes, and chatbot-based feedback mechanisms, helping students strengthen clinical reasoning and self-regulated learning [20]. These tools offer round-the-clock access to educational resources, addressing key challenges in medical and nursing education by providing scalable and adaptive learning solutions.

Despite their potential, the integration of chatbots into nursing education in China is still relatively limited [21]. Factors such as technological infrastructure, faculty readiness, and students' perceptions of chatbots play a significant role in determining how effectively they are adopted into curricula [22–24]. Cultural attitudes toward technology and education also shape how these tools are used in classrooms and clinical training environments [23]. Chatbots could transform the learning experience by providing instant feedback, personalized assistance, and easily accessible resources. Their success, however, relies heavily on the quality of their design and the way they are implemented within educational frameworks. While chatbots hold the promise of boosting student engagement and fostering higher-order thinking skills (HOTS)—such as critical thinking, problem-solving, and creativity [25–27]—realizing this potential requires careful study of the factors that influence their effectiveness.

The success of chatbot-based learning largely depends on two key factors: perceived usefulness (PUC) and ease of use (EoU), as outlined in the Technology Acceptance Model (TAM) [28]. Perceived usefulness reflects whether students believe chatbots can enhance their academic performance and help them achieve learning goals. At the same time, ease of use focuses on how intuitive and accessible chatbots are for users with varying technical skills. Beyond these core elements, feedback quality (FBQ) and self-regulated learning (SRL) act as mediating factors that transform engagement into meaningful learning outcomes. High-quality feedback ensures students receive timely, constructive responses to refine their understanding, while self-regulated learning encourages students to plan, monitor, and reflect on their educational journeys [29, 30]. Together, these factors create a comprehensive framework for understanding how chatbots can enhance learning experiences.

Each element plays a distinct but interconnected role in improving chatbot-based learning. Perceived usefulness motivates students to engage with chatbots they believe will help them succeed academically, such as by improving problem-solving or providing easy access to resources [31–34]. Ease of use reduces the frustration of poorly designed systems, ensuring students can navigate chatbot interfaces effortlessly [35, 36]. Engagement bridges usability and learning outcomes by involving students emotionally, behaviorally, and cognitively in their studies [37]. Feedback quality adds value to engagement by offering actionable, personalized insights tailored to individual learning needs [10, 38, 39]. Finally, self-regulated learning strengthens the link between engagement and higher-order thinking skills (HOTS) by promoting proactive and reflective learning behaviors essential for academic success [40, 41].

Despite their potential, chatbots in education still face several challenges. One major issue is the inconsistency of feedback quality. Many chatbots struggle to provide timely or nuanced responses, hindering students' ability to engage deeply with the content [42]. Poor interfaces can also deter students—especially those with limited technical skills—from using the tools regularly [10, 43]. Another hurdle is that not all students have strong self-regulation skills, limiting their ability to benefit from chatbot interactions fully [30]. Moreover, chatbots often fail to deliver the sophisticated support needed to develop higher-order cognitive skills like critical thinking, analysis, and creativity [44, 45].

To tackle these challenges, several solutions have been proposed. For instance, improving chatbot interfaces to be more user-friendly and intuitive helps increase accessibility and encourages consistent use [46, 47]. Advanced AI-driven personalization in feedback mechanisms ensures students receive tailored responses that align with their specific learning needs [25]. Chatbots can also promote self-regulated learning by providing structured prompts that guide students in planning, monitoring, and reflecting on their progress [48–50]. These measures aim to close the gap between student engagement and the development of higher-order thinking skills, ensuring every interaction with a chatbot contributes to meaningful learning outcomes.

However, some challenges remain. Customizing feedback to address the diverse needs of students and supporting the development of higher-order cognitive skills continues to be a struggle [29]. Scalability is

another issue, as implementing personalized features often requires advanced AI models that may not be feasible for large student populations [10, 51]. Additionally, while individual factors like PUC and EoU are well-documented, there is a lack of comprehensive models that integrate these elements and mediating factors like FBQ and SRL into a unified framework.

In summary, while significant progress has been made in addressing usability and engagement challenges, there is still much to learn about fully leveraging chatbots in education [17, 52, 53]. Combining perceived usefulness, ease of use, feedback quality, and self-regulated learning into a unified framework is essential to maximize the impact of chatbots on student engagement and learning outcomes. Bridging these gaps will pave the way for designing effective, user-friendly, and scalable chatbot systems that adapt to diverse educational contexts.

Literature Review

According to the Technology Acceptance Model (TAM) introduced by F Davis [28], individuals are more likely to adopt technology when they perceive it as beneficial to their performance. Chatbots are highly useful in education because they provide timely feedback, personalized learning experiences, and immediate support, fostering deeper student engagement [54–56]. Studies show that when students view chatbots as tools that improve outcomes like critical thinking and problem-solving, they are more willing to engage with them [2, 57, 58]. This is particularly true in higher education, where chatbots promote efficiency and self-directed learning [58]. Frequent interactions with chatbots perceived as helpful also lead to greater behavioral and cognitive engagement [24, 59].

The concept of perceived usefulness (PUC), introduced by F Davis [28] in the Technology Acceptance Model (TAM), posits that individuals are more likely to adopt and engage with a technology if they believe it enhances their performance. In the context of chatbots in education, perceived usefulness has been shown to influence student engagement significantly. Research indicates that students find chatbots useful because they provide timely feedback, personalized learning experiences, and immediate assistance, fostering deeper engagement [10].

Moreover, according to MA Almaiah, A Al-Khasawneh and A Althunibat [57], when students perceive chatbots as a tool for improving academic performance, such as critical thinking and problem-solving, their willingness to use them is boosted. In university-level educational environments, for example, chatbots make students' studies efficient and allow them to learn at their own pace [17]. Similarly, according to S Wollny, J Schneider, D Di Mitri, J Weidlich, M Rittberger and H Drachsler [58], students who perceived chatbots in a positive light and saw them as beneficial for them, then such students increased in terms of cognitive and behavioral engagement through direct use of chatbots at a high level.

Ease of use (EoU), another key TAM construct, refers to how effortless and intuitive technology is to use [28]. Research shows that user-friendly chatbot interfaces reduce cognitive load, allowing students to focus on learning rather than navigating complex systems [60, 61]. A Khamaj [47] found that intuitive chatbot designs significantly enhance engagement by encouraging repeated use and minimizing

frustration. Accessible interfaces are especially beneficial for students with limited technical skills, ensuring inclusivity and broadening engagement [10, 62]. In addition, in the context of chatbots, G-J Hwang, S-Y Wang and C-L Lai [63] observed that students who found the chatbot interface intuitive and easy to navigate were more likely to engage actively. User-friendly designs encourage repetitive use, fostering behavioral engagement while reducing frustration increases emotional engagement [17].

The interrelationship between ease of use and perceived usefulness is supported in educational technology research. V Venkatesh and FD Davis [64] extended the Technology Acceptance Model (TAM) and confirmed through experiments that ease of use indirectly impacts engagement (ENG) by enhancing perceived usefulness. For instance, when a chatbot is easy to use, students will perceive it as beneficial, leading to higher engagement. This indirect effect highlights the importance of designing chatbots that are both user-friendly and valuable in educational settings. O Zawacki-Richter, VI Marín, M Bond and F Gouverneur [25] reinforce this perspective, demonstrating that technologies perceived as both easy to use and useful promote higher cognitive and behavioral engagement. Specifically, chatbots that meet these criteria attract initial usage and sustain long-term engagement as students integrate them into their learning routines. Perceived ease of use and perceived usefulness are fundamental constructs that directly influence chatbot adoption in educational contexts. Existing research underscores the necessity of developing functional and intuitive chatbots to maximize their impact on student engagement. As academic institutions increasingly implement AI-driven technologies, understanding these relationships is crucial for enhancing learning outcomes and fostering meaningful interactions with digital tools.

High-quality feedback (FBQ) is essential for effective learning, as it helps students improve performance by clarifying goals and refining understanding [29]. Feedback quality bridges engagement and higher-order thinking skills (HOTS) in chatbot-based learning by transforming active participation into meaningful cognitive outcomes [65]. Timely, personalized, and constructive feedback has fostered skills like critical thinking and problem-solving [10, 25]. In chatbot-facilitated learning, feedback quality mediates the relationship between engagement and higher-order thinking skills by transforming active engagement into meaningful learning experiences [17]. For example, chatbots that provide tailored responses amplify the impact of engagement on cognitive skills, enabling students to analyse, evaluate, and create effectively [66].

Self-regulated learning (SRL) refers to learners' ability to plan, monitor, and evaluate their progress [30]. In chatbot-assisted environments, SRL mediates engagement and HOTS by channelling active participation into strategic learning behaviors [40, 41, 66]. Chatbots promote SRL by offering guidance and encouraging reflection. For instance, DH Chang, MP-C Lin, S Hajian and QQ Wang [67] observed that chatbots improve students' goal-setting and self-monitoring skills, essential for critical thinking and creativity. In chatbot-assisted learning environments, SRL acts as a mediator between engagement and higher-order thinking skills by channeling active participation into strategic learning behaviors [68]. Empirical studies have shown that chatbots promote SRL through reflection and scaffolding. For example, G-J Hwang, S-Y Wang and C-L Lai [63] observed that students who interacted with chatbots demonstrated improved self-regulation and goal-setting behaviors, which in turn enhanced their critical

thinking and creativity. Similarly, P Smutny and P Schreiberova [10] found that chatbots facilitate iterative learning processes, allowing students to develop and refine their problem-solving strategies and understanding over time.

Higher-order thinking skills (HOTS), which include analyzing, evaluating, and creating, are crucial for academic success and problem-solving [69]. These skills, situated at the top of Bloom's Taxonomy [70], are necessary for tackling complex problems and fostering innovation. Chatbots have shown significant potential in supporting HOTS by providing interactive and adaptive learning experiences [25].

Engagement is key to developing HOTS, as active participation encourages deeper cognitive processing. JA Fredricks, PC Blumenfeld and AH Paris [37] found that engaged students are likelier to analyse and synthesize information, core elements of critical thinking. Chatbots enhance these processes through timely feedback and scaffolding, which help students refine their understanding [29]. Moreover, creativity—an integral aspect of HOTS—is supported by chatbots through activities like brainstorming and open-ended prompts[26, 27]. Problem-solving skills are also nurtured as chatbots simulate real-world scenarios, guiding students through iterative processes to develop solutions [71].

While existing studies have explored the relationships between PUC, EoU, engagement, feedback quality, SRL, and HOTS, gaps remain in understanding how these elements interact within an integrated framework. Most research examines individual components, such as engagement or feedback, without fully addressing their combined or mediating roles in fostering HOTS. This highlights the need for comprehensive research that investigates these relationships holistically, including both direct and indirect effects. Such research could provide actionable insights for designing and optimizing chatbot-based learning environments to maximize their impact on student outcomes.

Theoretical Framework

This research draws on several theoretical frameworks to examine the relationships among perceived usefulness (PUC), ease of use (EoU), engagement, feedback quality, self-regulated learning, and higher-order thinking skills (HOTS). At its core is the Technology Acceptance Model (TAM), which emphasizes that perceived usefulness and ease of use drive technology adoption and sustained engagement [72]. F Davis [28] suggests that when students view chatbots as helpful and easy to use, their engagement increases, creating opportunities for transformative learning. Building on TAM, engagement theory highlights the importance of meaningful interaction, asserting that active cognitive, emotional, and behavioral involvement connects usability perceptions with mediators like feedback quality and self-regulated learning, ultimately influencing HOTS. J Hattie and H Timperley [29]' feedback model further explains how effective feedback enhances learning by clarifying goals, identifying areas for improvement, and encouraging deeper understanding through iterative processes. BJ Zimmerman [30]' self-regulated learning (SRL) theory adds another layer, emphasizing how planning, monitoring, and evaluating learning processes mediate the link between engagement and cognitive outcomes.

Lastly, Bloom's revised taxonomy [70] provides a framework for understanding the development of higher-order cognitive skills, such as analyzing, evaluating, and creating—outcomes that arise from meaningful engagement, high-quality feedback, and self-regulated learning. Together, these frameworks illuminate how chatbots can facilitate advanced learning processes and outcomes.

Hypothesis

Based on the above discussion, this study proposed the following hypothesis:

H1: Perceived usefulness of chatbots (PUC) has a positive direct effect on engagement (ENG).

H2: Ease of use of chatbots (EoU) has a positive direct effect on engagement (ENG).

H3: Engagement (ENG) has a positive direct effect on feedback quality (FBQ).

H4: Engagement (ENG) has a positive direct effect on self-regulated learning (SRL).

H5: Feedback quality (FBQ) mediates the relationship between engagement (ENG) and higher-order thinking skills (HOTS).

H6: Self-regulated learning (SRL) mediates the relationship between engagement (ENG) and higher-order thinking skills (HOTS).

The conceptual framework of this study was showed in Fig. 1.

Methods

Study Design

This study used a quantitative research design to explore the relationships among perceived usefulness (PUC), ease of use (EoU), engagement (ENG), feedback quality (FBQ), self-regulated learning (SRL), and higher-order thinking skills (HOTS). Data was collected through a cross-sectional survey, which allowed for the simultaneous measurement of these factors and the identification of their interrelationships. The study employed Structural Equation Modeling (SEM) to analyse the hypothesized relationships. This method would evaluate direct and indirect effects within the conceptual framework, providing a comprehensive understanding of how these constructs interact.

Participants and sample size

The target population for this study includes nursing students from various academic levels at the University X, which offers Nursing Education in China, all of whom have experience using educational chatbots as part of their coursework. Stratified random sampling was employed to ensure diverse

representation, with strata based on academic year (e.g., first-year, final-year students). This approach minimized sampling bias and captured differences in students' experiences and perspectives. A minimum sample size of 500 participants was determined following the recommendations for Structural Equation Modeling (SEM) [73], ensuring sufficient statistical power for detecting relationships and validating the model. A larger sample was targeted to account for non-responses or incomplete data. Participants were recruited through institutional email lists, learning management systems, and direct communication with faculty. Information about the study's purpose, confidentiality measures, and voluntary participation was provided to obtain informed consent. Participants were randomly selected within each stratum using computer-generated random numbers to ensure an unbiased selection process.

Instruments

A structured questionnaire was developed to measure six key constructs in the study: Perceived Usefulness (PUC), Ease of Use (EoU), Engagement (ENG), Feedback Quality (FBQ), Self-Regulated Learning (SRL), and Higher-Order Thinking Skills (HOTS). The items for PUC and EoU were adapted from F Davis [28] and V Venkatesh and FD Davis [64], while ENG was measured using the Student Engagement Scale [37]. FBQ items were custom-developed based on J Hattie and H Timperley [29]' feedback model, and SRL was assessed using BJ Zimmerman [30]' self-regulated learning framework. HOTS was measured with rubrics aligned with Bloom's Taxonomy [70], focusing on cognitive skills such as analysis and creation. All items were measured on a 7-point Likert scale (1 = Strongly Disagree to 7 = Strongly Agree) to ensure consistency and ease of response. The constructs capture specific aspects of the study, including system performance (PUC), usability (EoU), engagement levels (ENG), feedback clarity and relevance (FBQ), self-regulation strategies (SRL), and advanced cognitive skills (HOTS), ensuring a comprehensive assessment of the relationships under investigation. The summary of constructs, measurement items, and sources was shown in Table 1, For more details on the questionnaire, please see supplementary File 1.

Table 1
Summary of constructs, measurement items, and sources

Construct	Description	Source	Initial Items	Retained Items	Response Options
Perceived Usefulness (PUC)	Measures the extent to which users believe using the system improves their performance	F Davis [28]; V Venkatesh and FD Davis [64]	3	3	1 = Strongly Disagree, 7 = Strongly Agree
Ease of Use (EoU)	Assesses how simple and intuitive users find the system to operate.	F Davis [28]; V Venkatesh and FD Davis [64]	5	5	1 = Strongly Disagree, 7 = Strongly Agree
Engagement (ENG)	Evaluates the cognitive, emotional, and behavioral aspects of Engagement.	JA Fredricks, PC Blumenfeld and AH Paris [37]	5	5	1 = Strongly Disagree, 7 = Strongly Agree
Feedback Quality (FBQ)	Measures the clarity, relevance, and timeliness of feedback received.	J Hattie and H Timperley [29]; Custom items	4	4	1 = Strongly Disagree, 7 = Strongly Agree
Self-Regulated Learning (SRL)	Examines strategies like goal setting, self-monitoring, and self-reflection.	BJ Zimmerman [30]	5	5	1 = Strongly Disagree, 7 = Strongly Agree
Higher-Order Thinking Skills (HOTS)	Captures cognitive skills like analysis, evaluation, and creation.	LW Anderson and DR Krathwohl [70]	8	8	1 = Strongly Disagree, 7 = Strongly Agree

Validity and Reliability

To ensure content validity, a panel of experts from local nursing institutions and public universities specializing in educational technology and psychometrics reviewed the questionnaire. Construct validity will be assessed using Convergent Validity and Discriminant Validity via SmartPLS 4. In Partial Least Squares Structural Equation Modeling (PLS-SEM), construct validity is evaluated through specific tests rather than traditional Confirmatory Factor Analysis (CFA). Convergent validity will be determined by examining the Average Variance Extracted (AVE) for each construct; an AVE greater than 0.50 indicates that a construct explains over 50% of the variance in its indicators, which is considered acceptable [73]. Individual item factor loadings above 0.70 will also be deemed adequate. Reliability will be tested using Composite Reliability (CR) and Cronbach's Alpha, with CR values above 0.70 indicating internal consistency, supported further by Cronbach's Alpha [74]. Discriminant validity will be measured using the

Fornell-Larcker criterion. The Fornell-Larcker criterion requires that the square root of a construct's AVE be greater than its correlations with other constructs [74].

Data Collection and Analysis

Data was collected using an online survey platform to ensure accessibility and convenience for participants. The researchers sent 500 online survey forms to nursing students from years 1 to 4. A total of 480 forms were answered and submitted. However, 10 partially answered forms were excluded from the study. Finally, 470 participants were involved in this study, with a response rate of 94%. The data is exported to SPSS software to save. The analysis was conducted using Partial Least Squares Structural Equation Modeling (PLS-SEM) to test the hypothesised relationships, with the conceptual model evaluated through SmartPLS 4, a powerful tool for handling complex models with multiple constructs and maximizing explained variance [73]. The Standardized Root Mean Square Residual (SRMR) was reported to assess the overall model fit. SRMR is widely recognized as an appropriate fit index in PLS-SEM [74]. Although traditional fit indices like the Comparative Fit Index (CFI) and Root Mean Square Error of Approximation (RMSEA) are standard in covariance-based SEM, SRMR is more suitable for PLS-SEM and provides insights into the adequacy of the model.

Results

Table 2 presents an evaluation of the reliability and validity of the measurement model, confirming that all constructs meet or exceed recommended thresholds for internal consistency and convergent validity. Factor loadings range from 0.871 to 0.958, well above the 0.70 threshold [73], indicating strong item reliability. Composite Reliability (CR) values exceed 0.90 for all constructs, reflecting excellent internal consistency, with Higher-Order Thinking Skills (HOTS) achieving the highest CR value of 0.977. Cronbach's Alpha (α) values also demonstrate strong reliability, ranging from 0.924 to 0.973 across constructs. The Average Variance Extracted (AVE) values surpass the critical threshold of 0.50, with values ranging from 0.809 for Engagement (ENG) to 0.868 for Perceived Usefulness (PUC), confirming that the constructs explain a substantial proportion of variance in their indicators. Self-Regulated Learning (SRL) and Feedback Quality (FBQ) exhibit exceptionally high AVE values of 0.862 and 0.860, respectively, underscoring their decisive contributions to the model. These results confirm the measurement model's reliability, validity, and suitability for assessing the study's constructs.

Table 2
Loadings, composite reliability (CR), Cronbach's alpha, average variance extracted (AVE)

Construct	Loadings	Composite Reliability (CR)	Cronbach's Alpha (α)	Average Variance Extracted (AVE)
PUC		0.952	0.924	0.868
puc1	0.913			
puc2	0.958			
puc3	0.923			
EoU		0.964	0.953	0.842
eou1	0.913			
eou2	0.939			
eou3	0.918			
eou4	0.911			
eou5	0.908			
ENG		0.941	0.941	0.809
eng1	0.871			
eng2	0.882			
eng3	0.920			
eng4	0.926			
eng5	0.896			
SRL		0.969	0.960	0.862
srl1	0.940			
srl2	0.938			
srl3	0.934			
srl4	0.925			
srl5	0.906			
FBQ		0.961	0.945	0.860
fbq1	0.911			
fbq2	0.944			

Construct	Loadings	Composite Reliability (CR)	Cronbach's Alpha (α)	Average Variance Extracted (AVE)
fbq3	0.945			
fbq4	0.908			
HOTS		0.977	0.973	0.842
hots1	0.898			
hots2	0.918			
hots3	0.933			
hots4	0.923			
hots5	0.915			
hots6	0.925			
hots7	0.910			
hots8	0.917			

The Fornell-Larcker Criterion results (Table 3) confirm the discriminant validity of the measurement model, as the square root of the Average Variance Extracted (AVE) for each construct (diagonal values) is higher than its correlations with other constructs (off-diagonal values), in line with C Fornell and DF Larcker [75]' recommendations. For instance, the square root of AVE for Engagement (ENG) is 0.899, exceeding its highest correlation with Self-regulated Learning (SLR) at 0.892, indicating that ENG explains more variance in its indicators than in its relationships with other constructs. Similarly, the square root of AVE for Self-Regulated Learning (SRL) is 0.929, more significant than its correlation with Feedback Quality (FBQ) at 0.899, further supporting discriminant validity. The strongest correlations appear between SLR and FBQ (0.899) and FBQ and HOTS (0.912), reflecting their interdependence in the learning process. Meanwhile, Perceived Usefulness (PUC) shows weaker correlations with HOTS (0.687) and FBQ (0.701), suggesting it exerts an indirect effect. Overall, these findings validate the distinctiveness of the constructs while emphasizing the interconnected roles of engagement, feedback quality, and self-regulated learning in promoting higher-order thinking skills.

Table 3
Fornel -Larcker criterion

Constructs	ENG	EoU	FBQ	HOTS	PUC	SRL
ENG	0.899	0.804	0.852	0.840	0.711	0.892
EoU	0.804	0.918	0.773	0.749	0.829	0.833
FBQ	0.852	0.773	0.927	0.912	0.701	0.899
HOTS	0.840	0.749	0.912	0.918	0.687	0.872
PUC	0.711	0.829	0.701	0.687	0.932	0.737
SRL	0.892	0.833	0.899	0.872	0.737	0.929

Table 4 summarizes the f^2 analysis, highlighting the significant effects of various constructs within the model. Engagement (ENG) emerges as a central factor, exerting a very large effect on both Feedback Quality (FBQ) ($f^2 = 2.656$) and Self-Regulated Learning (SRL) ($f^2 = 3.874$), underscoring its crucial role in shaping feedback perceptions and promoting self-regulated learning behaviors. Ease of Use (EoU) shows a moderate to large effect on ENG ($f^2 = 0.424$), confirming that usability is vital in enhancing student engagement.

FBQ has a moderate to large effect on Higher-Order Thinking Skills (HOTS) ($f^2 = 0.470$), emphasizing the importance of high-quality feedback in fostering critical thinking. In contrast, $ENG \rightarrow HOTS$ ($f^2 = 0.030$) and $SRL \rightarrow HOTS$ ($f^2 = 0.025$) exhibit minor direct effects, suggesting that their impact on higher-order thinking is predominantly mediated through other pathways. Similarly, Perceived Usefulness (PUC) has a negligible effect on ENG ($f^2 = 0.019$), indicating that its influence may be more supportive than direct.

These results demonstrate the robustness of the model, with ENG serving as the central driver of significant outcomes. Supported by EoU, FBQ, and SRL, ENG optimizes learning outcomes and facilitates the development of critical thinking skills, highlighting its pivotal role in the learning process.

Table 4
f Square, effect size and interpretations

Predictor→ Outcome	f ² Value	Effect Size	Interpretation
ENG → FBQ	2.656	Very Large	Engagement has a powerful effect on Feedback Quality, emphasizing its critical role in improving feedback.
ENG → HOTS	0.030	Small	Engagement has a small direct effect on Higher-Order Thinking Skills, suggesting indirect mediation through other constructs like FBQ and SRL.
ENG → SRL	3.874	Very Large	Engagement has a powerful effect on Self-Regulated Learning, highlighting its importance in fostering self-regulation.
EoU → ENG	0.424	Medium to Large	Ease of Use has a moderate to significant effect on Engagement, indicating usability significantly influences engagement.
FBQ → HOTS	0.470	Medium to Large	Feedback Quality has a moderate to significant effect on Higher-Order Thinking Skills, demonstrating its role in fostering critical thinking.
PUC → ENG	0.019	Small	Perceived Usefulness has a negligible effect on Engagement, indicating a limited direct impact.
SRL → HOTS	0.025	Small	Self-Regulated Learning has a small direct effect on Higher-Order Thinking Skills, suggesting its influence is amplified through indirect pathways.

Table 5 shows the results from the SmartPLS bootstrapping analysis, which confirm that all hypothesized relationships are significant and support the proposed theoretical model. Perceived Usefulness (PUC) has a small but significant positive effect on Engagement (ENG) ($\beta = 0.143$, $p = 0.040$). In contrast, Ease of Use (EoU) has a strong positive effect on ENG ($\beta = 0.685$, $p = 0.000$), highlighting the critical roles of usability and perceived utility in driving user engagement. ENG significantly influences both Feedback Quality (FBQ) ($\beta = 0.852$, $p = 0.000$) and Self-Regulated Learning (SRL) ($\beta = 0.892$, $p = 0.000$), underscoring its pivotal role in enhancing feedback quality and fostering self-regulated learning behaviors. Mediation effects were also significant, with ENG → FBQ → HOTS showing a slight but notable effect ($\beta = 0.154$, $p = 0.019$) and ENG → SRL → HOTS demonstrating a slightly more substantial effect ($\beta = 0.169$, $p = 0.036$). These findings confirm the central role of engagement in improving learning outcomes both directly and indirectly, mediated through feedback quality and self-regulation learning. Overall, the results validate the theoretical framework and emphasize the interconnected roles of usability, engagement, feedback, and self-regulation learning in fostering higher-order thinking skills.

Table 5
Measurement model, p-value and decision

Hypothesis	Path	Path Coefficient (β)	p-value	Decision
H1	PUC $\rightarrow$ ENG (Perceived Usefulness $\rightarrow$ Engagement)	0.143	0.040	Supported (Significant)
H2	EoU $\rightarrow$ ENG (Ease of Use $\rightarrow$ Engagement)	0.685	0.000	Supported (Significant)
H3	ENG $\rightarrow$ FBQ (Engagement $\rightarrow$ Feedback Quality)	0.852	0.000	Supported (Significant)
H4	ENG $\rightarrow$ SRL (Engagement $\rightarrow$ Self-Regulated Learning)	0.892	0.000	Supported (Significant)
H5	ENG $\rightarrow$ FBQ $\rightarrow$ HOTS (Mediation Effect)	0.154	0.019	Supported (Significant)
H6	ENG $\rightarrow$ SRL $\rightarrow$ HOTS (Mediation Effect)	0.169	0.036	Supported (Significant)

Figure 2 illustrates the structural model and the relationships among Perceived Usefulness (PUC), Ease of Use (EoU), Engagement (ENG), Feedback Quality (FBQ), Self-Regulated Learning (SRL), and Higher-Order Thinking Skills (HOTS). Path coefficients (β) reveal significant relationships, with EoU $\rightarrow$ ENG ($\beta = 0.685$) showing a strong effect, while PUC $\rightarrow$ ENG ($\beta = 0.143$) has a more minor but significant impact. Engagement plays a central role, exerting substantial effects on FBQ ($\beta = 0.852$) and SRL ($\beta = 0.892$), mediating its influence on HOTS. Indirect effects through FBQ ($\beta = 0.154$) and SRL ($\beta = 0.169$) amplify the total effect of ENG on HOTS ($\beta = 0.952$), confirming partial mediation. The R^2 values reflect strong explanatory power, with HOTS ($R^2 = 0.850$) primarily explained by ENG, FBQ, and SRL, while ENG significantly influences FBQ ($R^2 = 0.727$) and SRL ($R^2 = 0.795$). Indicator loadings, such as PUC1 = 0.913 and ENG3 = 0.920, exceed the 0.70 threshold, confirming the reliability of the measurement model. The figure highlights the pivotal role of ENG in driving learning outcomes, with FBQ and SRL amplifying its impact on HOTS. These findings validate the model's strong predictive power and alignment with the theoretical framework.

Table 6 presents the mediation analysis, confirming that the relationship between Engagement (ENG) and Higher-Order Thinking Skills (HOTS) is partially mediated by both Feedback Quality (FBQ) and Self-Regulated Learning (SRL). The direct effect of ENG on HOTS is $\beta = 0.154$ ($p = 0.019$), while the indirect effect through FBQ and SRL is substantial ($\beta = 0.798$), resulting in a total effect of $\beta = 0.952$. This indicates that while ENG has a direct influence on HOTS, a significant portion of its impact is mediated by these constructs, confirming partial mediation.

Specifically, the pathway ENG $\rightarrow$ FBQ $\rightarrow$ HOTS has an indirect effect of $\beta = 0.154$, highlighting the significant role of feedback quality in mediating engagement's effect on higher-order thinking. Similarly, the pathway ENG $\rightarrow$ SRL $\rightarrow$ HOTS exhibits a slightly stronger indirect effect of $\beta = 0.169$, emphasizing

the critical role of self-regulated learning as a mediator. Between the two mediators, SRL demonstrates a slightly greater influence than FBQ.

The analysis also reveals full mediation for predictors like Perceived Usefulness (PUC) and Ease of Use (EoU) on HOTS. The pathway PUC → ENG → HOTS shows an indirect effect of $\beta = 0.022$, while EoU → ENG → HOTS has a stronger indirect effect of $\beta = 0.106$, both entirely mediated through engagement.

These findings underscore the pivotal roles of FBQ and SRL in enhancing HOTS, while constructs like PUC and EoU influence HOTS indirectly through engagement. The results validate the interconnected pathways in the model, highlighting the importance of feedback quality and self-regulation in fostering cognitive outcomes. Additionally, model fit indices were evaluated to assess how well the structural equation model (SEM) aligns with the observed data, ensuring the relationships specified in the model accurately reflect the theoretical framework.

Table 6
Mediation analysis matrix

Path	Direct Effect (β)	Indirect Effect (β)	Total Effect (β)	Decision
ENG → HOTS	0.154	0.798 (via FBQ & SRL)	0.952	Partial Mediation
ENG → FBQ → HOTS	-	0.154	0.154	Supported (Mediation)
ENG → SRL → HOTS	-	0.169	0.169	Supported (Mediation)
PUC → ENG → HOTS	-	$0.143 \times 0.154 = 0.022$	0.022	Full Mediation
EoU → ENG → HOTS	-	$0.685 \times 0.154 = 0.106$	0.106	Full Mediation
ENG → FBQ	0.852	-	0.852	Direct Effect
ENG → SRL	0.892	-	0.892	Direct Effect
FBQ → HOTS	0.629	-	0.629	Direct Effect
SRL → HOTS	0.169	-	0.169	Direct Effect

Table 7 compares the model fit indices for the saturated (perfectly fitting) model and the estimated model. The SRMR (Standardized Root Mean Square Residual) for the estimated model is 0.063, within the acceptable range (below 0.08), indicating a reasonable fit, though slightly worse than the saturated model (0.033). The d_ULS (1.841) and d_G (0.967) values for the estimated model are higher than those of the saturated model (0.498 and 0.809, respectively), suggesting minor model misspecifications.

The Chi-square value for the estimated model (2345.283) is also higher than that of the saturated model (2157.606), reflecting some deviation between the observed and model-implied covariance matrices. Lastly, the NFI (Normed Fit Index) for the estimated model is 0.883, slightly below the commonly accepted threshold of 0.90. While these results indicate that the model has an acceptable fit, they also highlight areas where refinement could improve its alignment with the data.

Table 7
Model fit indices: comparison of saturated and estimated models

Fit Index	Saturated Model	Estimated Model
SRMR	0.033	0.063
d_ULS	0.498	1.841
d_G	0.809	0.967
Chi-Square	2157.606	2345.283
NFI	0.893	0.883

Discussion

This study highlights the transformative potential of thoughtfully designed AI-powered chatbots in enhancing nursing education, particularly by fostering engagement and higher-order thinking skills (HOTS). Consistent with the Technology Acceptance Model (TAM), Perceived Usefulness (PUC) and Ease of Use (EoU) emerged as crucial drivers of Engagement (ENG). Nursing students who found chatbots intuitive and beneficial demonstrated greater motivation and active participation in their learning, underscoring the importance of user-centric designs tailored to meet the unique demands of nursing education, such as intuitive interfaces and practical clinical applications.

Aligned with TAM [28, 64], the positive effects of PUC and EoU on engagement reflect how intuitive and functional chatbot designs can support nursing students in managing high academic demands. Chatbots that combine ease of use with practical functionality make learning more efficient and accessible, emphasizing the need for interfaces designed specifically for the challenges of nursing education.

While engagement strongly influences Feedback Quality (FBQ) and Self-Regulated Learning (SRL), its direct effect on HOTS is minimal. This finding confirms that engagement alone cannot drive deep learning but must be paired with meaningful strategies to achieve cognitive and professional growth [37, 76–78]. Chatbots should therefore go beyond simple interactions by offering personalized feedback and scaffolds that support critical thinking, clinical decision-making, and reflective practices essential for professional competency.

Both FBQ and SRL are vital mediators that transform engagement into HOTS. High-quality feedback—timely, clear, and relevant—helps nursing students critically analyze their performance and refine essential clinical skills [29, 79]. Similarly, SRL equips students with critical skills like goal-setting, self-monitoring, and reflection [30, 41]. These findings highlight the importance of chatbots that foster autonomous learning and provide structured, adaptive support for nursing students navigating the rigorous demands of their education.

Despite their promise, AI-powered chatbots in nursing education face challenges such as inconsistent feedback quality and variability in students' self-regulation abilities [10, 42]. Nursing students in China, often constrained by limited resources, limited access to clinical training, and scalability issues[22], stand to benefit from innovations like adaptive feedback algorithms and personalized learning paths[22, 25]. Scalable, culturally tailored solutions are key to maximizing the impact of chatbot-based learning.

Limitations

This study has several limitations. First, its cross-sectional design limits the ability to infer causality. Second, all data were self-reported, which may introduce response bias. Third, although stratified sampling was employed, the sample was limited to students from one university in China, which may affect the generalizability of the results. Finally, the study relied solely on quantitative data; future research could benefit from mixed-method approaches to gain deeper insights into students' experiences with chatbot-based learning.

Conclusions

AI-powered chatbots hold significant promise as virtual learning assistants for nursing students in China. Their effectiveness, however, is influenced by several factors, including design quality, implementation context, and individual learner characteristics. This study highlights the pivotal roles of PUC and EoU in fostering student engagement. In turn, engagement enhances FBQ and SRL, both of which serve as key mediators in the development of HOTS.

Nevertheless, engagement alone is not sufficient to promote deep learning. To maximize their educational impact, chatbots must be designed to deliver timely, personalized feedback and effectively support self-regulation [29, 76, 77]. Addressing persistent challenges such as feedback variability and limited scalability is critical to optimizing chatbot-based learning environments.

Future research should prioritize the development of advanced feedback algorithms, ensure alignment of chatbot functions with authentic clinical scenarios, and incorporate longitudinal and experimental designs to assess their long-term impact on professional competence and patient care outcomes. By overcoming current limitations and leveraging AI advancements, chatbots have the potential to transform nursing education, equipping students with the critical thinking and problem-solving skills needed to thrive in an evolving healthcare landscape.

Declarations

Acknowledgements

The authors would like to thank all nursing students who participated in this study, Mr. Huo Xincun, Ms. Zhai Shuna and all of the experts who provided validation, support, and recommendations.

Author's contributions

RW: Conceptualization, Methodology, Investigation, Data Curation, Writing, Original Draft, Writing - Review & Editing, Supervision, Project administration. AR: Conceptualization, Methodology, Validation, Data Curation, Writing, Original Draft, Writing - Review & Editing, Supervision, Project administration. All authors read and approved the final manuscript.

Funding

This research was supported by the Shandong Province Undergraduate Teaching Reform Project (M2022022).

Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

The ethical clearance for this study was obtained from Shandong Xiehe University's Ethics Committee (reference LLSC-KY03-2025004). The study was conducted in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. Information about the study's purpose, confidentiality measures, and voluntary participation was provided, and informed consent was obtained from all participants.

Clinical trial number: not applicable

Consent for publication

Not applicable.

Competing Interests

The authors declare that they have no competing interests.

References

1. Ma YZ, Wang J, Li X, Sun W, Wu H, Liu T, Zhao QH, Xiao MC. Qualitative study on nursing students' experience in the use of large language model chatbots. *J Nurs Sci*. 2024;39(16):69–72.
2. Labadze L, Grigolia M, Machaidze L. Role of AI chatbots in education: systematic literature review. *Int J Educational Technol High Educ*. 2023;20(1):56.
3. Kooli C. Chatbots in Education and Research: A Critical Examination of Ethical Implications and Solutions. *Sustainability*. 2023;15(7):5614.
4. Tam W, Huynh T, Tang A, Luong S, Khatri Y, Zhou W. Nursing education in the age of artificial intelligence powered Chatbots (AI-Chatbots): Are we ready yet? *Nurse Educ Today*. 2023;129:105917.
5. Chen X, Qian X. Overview of Healthcare System in China. In: *Quality Spine Care: Healthcare Systems, Quality Reporting, and Risk Adjustment*. edn. Edited by Ratliff J, Albert TJ, Cheng J, Knightly J. Cham: Springer International Publishing; 2019: 237–254.
6. Chen C, Liu M. Achievements and Challenges of the Healthcare System in China. *Cureus*. 2023;15(5):e39030.
7. Chang CY, Kuo SY, Hwang GH. Chatbot-facilitated Nursing Education Incorporating a Knowledge-Based Chatbot System into a Nursing Training Program. *Educational Technol Soc*. 2022;25(1):15–27.
8. Wang CQ, Xiao HL, Wu YX, Li WJ, Wang XY, Liu H, Xiao Y, Xing XY, Hu WY. Application and perspectives of large language model in the field of nursing: Represented by ChatGPT. *J Nurses Train*. 2024;39(12):1296–300.
9. Kuhail MA, Alturki N, Alramlawi S, Alhejori K. Interacting with educational chatbots: A systematic review. *Educ Inform Technol*. 2023;28(1):973–1018.
10. Smutny P, Schreiberova P. Chatbots for learning: A review of educational chatbots for the Facebook Messenger. *Comput Educ*. 2020;151:103862.
11. Srinivasan M, Venugopal A, Venkatesan L, Kumar R. Navigating the Pedagogical Landscape: Exploring the Implications of AI and Chatbots in Nursing Education. *JMIR Nurs*. 2024;7:e52105.
12. Pérez JQ, Daradoumis T, Puig JMM. Rediscovering the use of chatbots in education: A systematic literature review. *Comput Appl Eng Educ*. 2020;28(6):1549–65.
13. Qu X, Yang JM, Chen T, Zhang W. Reflections on the Implications of the Developments in ChatGPT for Changes in Medical Education Models. *J Sichuan Univ (Medical Sciences)*. 2023;54(5):937–40.
14. Labrague L, Al Sabei S. Integration of AI-Powered Chatbots in Nursing Education: A Scoping Review of Their Utilization, Outcomes, and Challenges. *Teach Learn Nurs* 2024, 20.
15. Rodriguez-Arrastia M, Martinez-Ortigosa A, Ruiz-Gonzalez C, Ropero-Padilla C, Roman P, Sanchez-Labraca N. Experiences and perceptions of final-year nursing students of using a chatbot in a simulated emergency situation: A qualitative study. *J Nurs Manag*. 2022;30(8):3874–84.
16. Chang C-Y, Yang C-L, Jen H-J, Ogata H, Hwang G-H. Facilitating nursing and health education by incorporating ChatGPT into learning designs. *Educational Technol Soc*. 2023;27(1):215–30.

17. Zhang F, Liu X, Wu W, Zhu S. Evolution of Chatbots in Nursing Education: Narrative Review. *JMIR Med Educ.* 2024;10:e54987.
18. Teachers' pet. iFlytek's AI-powered education technology
[<https://www.thewirechina.com/2024/06/09/teachers-pet-iflytek-ai-edtech-education-technology/>]
19. Chinese AI. 'Xiaoyi' passes national medical licensing test
[<https://www.therobotreport.com/chinese-xiaoyi-ai-platform-passes-national-medical-licensing-tests/>]
20. Okonkwo CW, Ade-Ibijola A. Chatbots applications in education: A systematic review. *Computers Education: Artif Intell.* 2021;2:100033.
21. Jiang Y, Kong M. The evolution of artificial intelligence on nursing education in China. *Multidisciplinary Reviews* 2024, 7(12).
22. Jiang Y. Artificial Intelligence in Nursing Education: Challenges and Opportunities in the Chinese Context. *Open J Social Sci.* 2024;12(11):55–66.
23. Stöhr C, Ou AW, Malmström H. Perceptions and usage of AI chatbots among students in higher education across genders, academic levels and fields of study. *Computers Education: Artif Intell.* 2024;7:100259.
24. Ayanwale MA, Ndlovu M. Investigating factors of students' behavioral intentions to adopt chatbot technologies in higher education: Perspective from expanded diffusion theory of innovation. *Computers Hum Behav Rep.* 2024;14:100396.
25. Zawacki-Richter O, Marín VI, Bond M, Gouverneur F. Systematic review of research on artificial intelligence applications in higher education – where are the educators? *Int J Educational Technol High Educ.* 2019;16(1):39.
26. Karanjakwut C, Charunsri K. Transforming AI Chatbots for a Brainstorming Teaching Technique of Process Writing. *Malaysian Online J Educational Technol.* 2025;13:1–18.
27. Moretti D. Nowadays many brainstorming idea with chatbots; 2024.
28. Davis F. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Q.* 1989;13:319.
29. Hattie J, Timperley H. The Power of Feedback. *Rev Educ Res.* 2007;77(1):81–112.
30. Zimmerman BJ. Becoming a Self-Regulated Learner: An Overview. *Theory Into Pract.* 2002;41(2):64–70.
31. Chen H-L, Vicki Widarso G, Sutrisno H. A ChatBot for Learning Chinese: Learning Achievement and Technology Acceptance. *J Educational Comput Res.* 2020;58(6):1161–89.
32. Liao Y-K, Wu W-Y, Le TQ, Phung TTT. The Integration of the Technology Acceptance Model and Value-Based Adoption Model to Study the Adoption of E-Learning: The Moderating Role of e-WOM. *Sustainability* 2022, 14(2).
33. Abouammoh N, Alhasan K, Raina R, Malki K, Aljamaan F, Tamimi I, Muaygil R, Wahabi H, Jamal A, Al-Tawfiq J et al. Exploring Perceptions and Experiences of ChatGPT in Medical Education: A

- Qualitative Study Among Medical College Faculty and Students in Saudi Arabia. *medRxiv* 2023.
34. Bonsu E, Baffour-Koduah D. From the Consumers' Side: Determining Students' Perception and Intention to Use ChatGPT in Ghanaian Higher Education, vol. 4; 2023.
 35. Chan CKY, Hu W. Students' voices on generative AI: perceptions, benefits, and challenges in higher education. *Int J Educational Technol High Educ*. 2023;20(1):43.
 36. Ali J, Shamsan M, Hezam T, Majeed A. Impact of ChatGPT on Learning Motivation: Teachers and Students' Voices. *J Engl Stud Arabia Felix*. 2023;2:41–9.
 37. Fredricks JA, Blumenfeld PC, Paris AH. School Engagement: Potential of the Concept, State of the Evidence. *Rev Educ Res*. 2004;74(1):59–109.
 38. Guo K, Pan M, Li Y, Lai C. Effects of an AI-supported approach to peer feedback on university EFL students' feedback quality and writing ability. *Internet High Educ*. 2024;63:100962.
 39. Guo K, Wang D. To resist it or to embrace it? Examining ChatGPT's potential to support teacher feedback in EFL writing. *Educ Inform Technol*. 2024;29(7):8435–63.
 40. Lee H-Y, Chen P-H, Wang W-S, Huang Y-M, Wu T-T. Empowering ChatGPT with guidance mechanism in blended learning: effect of self-regulated learning, higher-order thinking skills, and knowledge construction. *Int J Educational Technol High Educ*. 2024;21(1):16.
 41. Azevedo R, Gašević D. Analyzing Multimodal Multichannel Data about Self-Regulated Learning with Advanced Learning Technologies: Issues and Challenges. *Comput Hum Behav*. 2019;96:207–10.
 42. Rhim J, Kwak M, Gong Y, Gweon G. Application of humanization to survey chatbots: Change in chatbot perception, interaction experience, and survey data quality. *Comput Hum Behav*. 2022;126:107034.
 43. Roldán-Álvarez D, Martín E, García-Herranz M, Haya PA. Mind the gap: Impact on learnability of user interface design of authoring tools for teachers. *Int J Hum Comput Stud*. 2016;94:18–34.
 44. Zhai C, Wibowo S, Li LD. The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review. *Smart Learn Environ*. 2024;11(1):28.
 45. Krupp L, Steinert S, Kiefer-Emmanouilidis M, Avila K, Lukowicz P, Kuhn J, Küchemann S, Karolus J. Unreflected Acceptance – Investigating the Negative Consequences of ChatGPT-Assisted Problem Solving in Physics Education. In: *HHAI 2024: Hybrid Human AI Systems for the Social Good*. edn.; 2024.
 46. Best Practices for Designing User-Friendly AI Chatbot Interfaces.
[<https://www.linkedin.com/pulse/best-practices-designing-user-friendly-ai-chatbot-interfaces-9m4kc>]
 47. Khamaj A. AI-enhanced chatbot for improving healthcare usability and accessibility for older adults. *Alexandria Eng J*. 2025;116:202–13.
 48. Guan R, Raković M, Chen G, Gašević D. How educational chatbots support self-regulated learning? A systematic review of the literature. *Educ Inform Technol* 2024.

49. Yin J, Goh T-T, Yang B, Xiaobin Y. Conversation Technology With Micro-Learning: The Impact of Chatbot-Based Learning on Students' Learning Motivation and Performance. *J Educational Comput Res.* 2021;59(1):154–77.
50. Hew K, Huang W, Du J, Jia C. Using Chatbots in Flipped Learning Online Sessions: Perceived Usefulness and Ease of Use. In., edn.: Springer, Cham;; 2021. pp. 164–75.
51. Kamalov F, Santandreu Calonge D, Gurrib I. New Era of Artificial Intelligence in Education: Towards a Sustainable Multifaceted Revolution. *Sustainability* 2023, 15(16).
52. Choi EPH, Lee JJ, Ho MH, Kwok JYY, Lok KYW. Chatting or cheating? The impacts of ChatGPT and other artificial intelligence language models on nurse education. *Nurse Educ Today.* 2023;125:105796.
53. Jallad ST, Alsaqer K, Albadareen BI, Al-maghaireh D. Artificial intelligence tools utilized in nursing education: Incidence and associated factors. *Nurse Educ Today.* 2024;142:106355.
54. Molnár G, Szüts Z. The role of chatbots in formal education. In: *2018 IEEE 16th International Symposium on Intelligent Systems and Informatics (SISY): 2018; Subotica, Serbia: IEEE; 2018: 000197–000202.*
55. Lam C, Chan L, See C: Converse, connect and consolidate–The development of an artificial intelligence chatbot for health sciences education. In: *Frontiers in Medical and Health Sciences Education 2018: Learning in Alliance: Inter-professional Health Education and Practice: 2018; Hong Kong: Bau Institute of Medical and Health Sciences Education, Li Ka Shing Faculty …*
56. Adamopoulou E, Moussiades L. Chatbots: History, technology, and applications. *Mach Learn Appl.* 2020;2:100006.
57. Almaiah MA, Al-Khasawneh A, Althunibat A. Exploring the critical challenges and factors influencing the E-learning system usage during COVID-19 pandemic. *Educ Inform Technol.* 2020;25(6):5261–80.
58. Wollny S, Schneider J, Di Mitri D, Weidlich J, Rittberger M, Drachsler H. Are We There Yet? - A Systematic Literature Review on Chatbots in Education. *Front Artif Intell.* 2021;4:654924.
59. Goli M, Sahu DA, Bag S, Dhamija P. Users' Acceptance of Artificial Intelligence-Based Chatbots: An Empirical Study. *Int J Technol Human Interact.* 2023;19:1–18.
60. Sweller J. Cognitive Load During Problem Solving: Effects on Learning. *Cogn Sci.* 1988;12(2):257–85.
61. Reis HM, Borges SS, Durelli VHS, Fernando de S, Moro L, Brandao AAF, Barbosa EF, Brandao LO, Isotani S, Jaques PA, Bittencourt II. Towards Reducing Cognitive Load and Enhancing Usability through a Reduced Graphical User Interface for a Dynamic Geometry System: An Experimental Study. In: *2012 IEEE International Symposium on Multimedia.* 2012: 445–450.
62. Pandita A, Kiran R. The Technology Interface and Student Engagement Are Significant Stimuli in Sustainable Student Satisfaction. *Sustainability.* 2023;15:7923.
63. Hwang G-J, Wang S-Y, Lai C-L. Effects of a social regulation-based online learning framework on students' learning achievements and behaviors in mathematics. *Comput Educ.* 2021;160:104031.

64. Venkatesh V, Davis FD. A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. *Manage Sci.* 2000;46(2):186–204.
65. Fadillah MA, Usmeldi U, Asrizal A. The Role of Chatgpt and Higher-Order Thinking Skills as Predictors of Physics Inquiry. *J Baltic Sci Educ.* 2024;23(6):1178–92.
66. Lin MP-C, Chang D. CHAT-ACTS: A pedagogical framework for personalized chatbot to enhance active learning and self-regulated learning. *Computers Education: Artif Intell.* 2023;5:100167.
67. Chang DH, Lin MP-C, Hajian S, Wang QQ. Educational Design Principles of Using AI Chatbot That Supports Self-Regulated Learning in Education: Goal Setting, Feedback, and Personalization. *Sustainability* 2023, 15(17).
68. Taub M, Sawyer R, Smith A, Rowe J, Azevedo R, Lester J. The agency effect: The impact of student agency on learning, emotions, and problem-solving behaviors in a game-based learning environment. *Comput Educ.* 2020;147:103781.
69. Widana IW, HIGHER ORDER THINKING SKILLS ASSESSMENT (HOTS). *JISAE: J Indonesian Student Assess Evaluation.* 2017;3:32–44.
70. Anderson LW, Krathwohl DR. *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives.* Longman; 2001.
71. Ali M, Guibas L. *ChatGPT and Cognitive Skills Development: Integrating Generative AI in Modern Pedagogics Design;* 2024.
72. Kryvinska MJPVs. Exploring the Chatbot usage intention-A mediating role of Chatbot initial trust. *Heliyon.* 2024;10(12):e33028.
73. Hair J, Hult GTM, Ringle C, Sarstedt M. *A Primer on Partial Least Squares Structural Equation Modeling. PLS-SEM);* 2022.
74. Henseler J, Ringle CM, Sarstedt M. A new criterion for assessing discriminant validity in variance-based structural equation modeling. *J Acad Mark Sci.* 2015;43(1):115–35.
75. Fornell C, Larcker DF. Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *J Mark Res.* 1981;18(1):39–50.
76. Mystakidis S. Deep Meaningful Learning. *Encyclopedia.* 2021;1:988–97.
77. Bhardwaj P, Gupta PK, Panwar H, Siddiqui MK, Morales-Menendez R, Bhaik A. Application of Deep Learning on Student Engagement in e-learning environments. *Comput Electr Eng.* 2021;93:107277.
78. Alabi M. *Experiential Learning: Fostering Deep Learning Through Active Engagement.* 2024.
79. Burgess A, van Diggele C, Roberts C, Mellis C. Feedback in the clinical setting. *BMC Med Educ.* 2020;20(2):460.

Figures

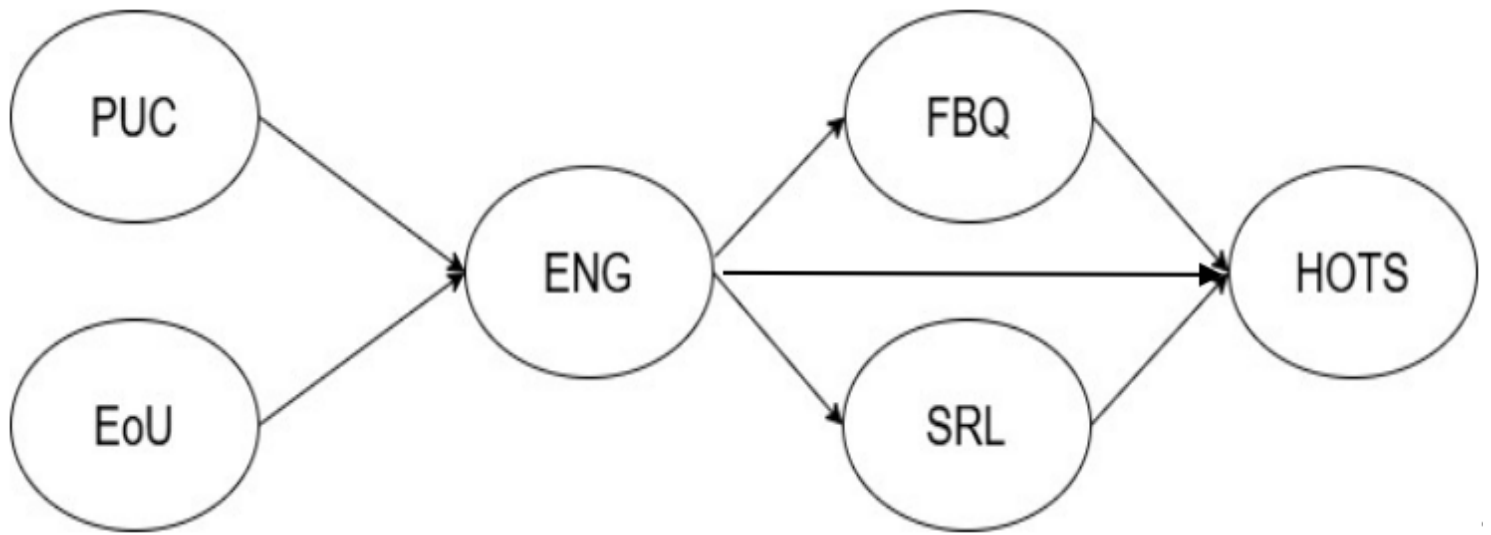


Figure 1

Conceptual framework of the study

Note: PUC- Perceived usefulness of chatbots, EoU- Ease of use of chatbots, ENG- Engagement with chatbots, FBQ- Feedback quality, SRL- Self-regulated learning, HOTS- Higher-order thinking skills

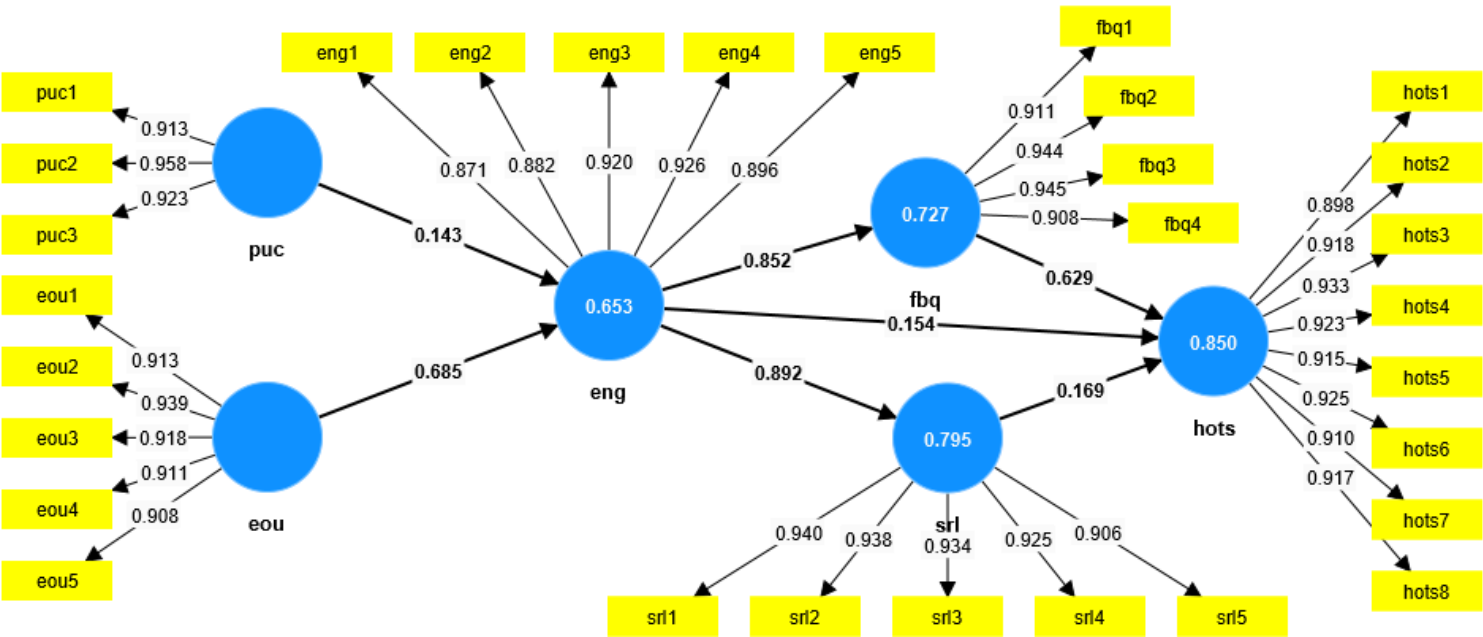


Figure 2

Structural Equation Model

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementary1ResearchQuestionnaire.docx](#)