Han et al. BMC Medical Education (2025) 25:825
https://doi.org/10.1186/s12909-025-07316-2

BMC Medical Education

RESEARCH

Open Access

Development and effects of a chatbot education program for self-directed learning in nursing students

Jeong-won Han[1], Junhee Park[2] and Hanna Lee[3*]

Abstract

Background Critical care nurses require adequate thinking skills to assess the condition of acutely-ill patients, make decisions, and implement appropriate interventions. Specifically, precise knowledge and swift response to mechanical ventilation-related nursing are of utmost importance. Educational needs of clinical nurses revealed a high demand for education in mechanical ventilation nursing. This study aims to develop a chatbot educational program for self-directed learning to enhance nursing students' skills and assess its impact on clinical reasoning competency, mechanical ventilation nursing-related knowledge, self-confidence, and education satisfaction.

Methods This study employed a randomized controlled trial with a pretest-posttest parallel group design. The participants included fourth-year nursing students from the Department of Nursing at a university in the "G" province in South Korea, with 31 participants in the experimental group and 29 in the control group. The experimental group had access to the Uniform Resource Locator or URL of the mechanical ventilation nursing chatbot educational program and video lectures, while the control group received only the video lectures.

Results The experimental group demonstrated statistically significantly higher clinical reasoning competency (t=-5.00, p <.001), self-confidence (t=-2.62, p =.011), and satisfaction in education (t=-3.51, p <.001) compared to the control group. However, there was no statistically significant difference in knowledge (t=-0.09, p =.926).

Conclusions The chatbot educational program could be an effective educational method to positively influence nursing students' clinical reasoning competency, self-confidence, and education satisfaction. It can enhance self-directed learning among nursing students and effectively improve nursing competency.

Clinical trial number Not applicable.

Keywords Clinical reasoning, Education, Nurse, Student, Knowledge, Mechanical ventilation

*Correspondence:
Hanna Lee
hannalee@gwnu.ac.kr
[1]College of Nursing Science, Kyung Hee University, Seoul, Republic of Korea
[2]Department of Nursing, Dongnam Health University, Suwon, Gyeonggi-do, Republic of Korea
[3]Department of Nursing, Gangneung-Wonju National University, Wonju-si, Republic of Korea

Background

Critical care nursing is a specialized field that involves caring for acutely-ill patients who are unstable and prone to rapid deterioration owing to actual or potential life-threatening health issues. Critical care nurses require adequate thinking skills to assess the condition of acutely-ill patients, make decisions, and implement appropriate interventions [1]. A significant proportion of patients that are admitted to the intensive care unit exhibit signs of respiratory failure, making the care of mechanically-ventilated patients a regular task of nurses responsible for managing critically-ill patients. This care is vital and directly linked to a patient's life. Inappropriate management of mechanical ventilation, such as ventilator discontinuation, poses a threat to patients, necessitating nurses' vigilant monitoring and accurate decision-making in response to rapidly-changing patient conditions [2]. Therefore, it should be emphasized that precise knowledge and swift response to mechanical ventilation-related nursing are of utmost importance for nurses [3].

A study of the educational needs of clinical nurses determined that there is a high demand for education in mechanical ventilation nursing [4], which includes specialized knowledge and advanced skills, necessitating repeated education and training for clinical nurses. Additionally, it is essential for nursing students who are aspiring healthcare professionals to learn and acquire skills related to mechanical ventilation nursing [5]. In the wake of the spread of infectious diseases such as COVID-19, there is growing societal demand for skilled nursing professionals capable of providing critical care. However, clinical training for students is severely limited amid concerns of patient safety and other issues [6]. Nevertheless, the restricted availability of critical care nursing practicum for students has raised concerns among educators aiming to produce specialized critical care nursing professionals [6].

Mechanical ventilation nursing education for nursing students has been primarily relying on methods such as lectures or high-fidelity simulation (HFS) [7]. Simulation education offers the advantage of experiencing interactive scenarios with patients and healthcare professionals based on real-life situations [8]. However, significant costs are involved in the installation and maintenance of simulation equipment. Furthermore, in cases where group education in specific locations is restricted owing to infectious diseases, there are limitations in conducting programs using standardized patients and simulators [9]. Accordingly, multimedia education is being implemented as a cost-effective alternative to HFS with advantages such as improved accessibility, flexibility, learner-centeredness, and expansibility [10]. However, there are limitations to experiencing repetitive self-directed learning (SDL) through this approach. Ultimately, in cases requiring repetitive education, such as mechanical ventilation nursing, it is essential to create an environment where learners can engage in self-study and address learning areas that they may be lacking.

In SDL, individuals control learning by assuming a strong sense of self-responsibility. They regulate their learning continually and assess and monitor their learning progress while maintaining autonomy [11]. This approach is gaining attention as a method to enhance academic achievement in educational settings. SDL encompasses learners' ability to plan learning, regulate the learning environment, adjust cognition and self-monitoring methods for learning tasks (self-monitoring method), and employ cognitive strategies to remember and comprehend the learning material [12]. SDL can be seen as a learner's ability to assess and control their learning status, set learning goals, adjust learning pace, and continually check and evaluate whether the goals they aim to achieve during the learning process are being met [13]. Therefore, there is a growing need to focus on chatbot programs to facilitate SDL.

A chatbot is a type of software that can engage in conversations with humans through speech or text, perform user requests, and engage in everyday dialogues in a question-and-answer format [14]. Chatbots have been employed in educational programs to engage students in interactive learning, fostering knowledge construction from a constructivist perspective. Extensive efforts have been made to enhance nursing education through programs based on artificial intelligence (AI) that provide personalized interactions [15]. Chatbots, free from time and space constraints, can complement the limitations of simulation education, enabling learners to engage in SDL by facilitating conversations anytime and anywhere [16]. Furthermore, chatbots can enhance clinical reasoning competency among nursing students by providing various materials and scenarios, allowing them to experience clinical judgments [17].

Nursing research has been predominantly focusing on the utilization of chatbots in areas such as mental health [18, 19], weight management [20, 21], antenatal education [22], and electronic fetal monitoring [16]. However, there is a need for research related to mechanical ventilation. In the wake of advanced knowledge, skills, clinical judgment abilities, and rapidly-evolving clinical evidence, it is crucial to utilize AI in educational programs to enable nursing students to learn and practice safely and efficiently, especially in areas like mechanical ventilation [23]. Additionally, research conducted in Hong Kong on a chatbot educational program implemented for graduate students showed that learners benefited from increased motivation, satisfaction, and knowledge improvement through chatbot-assisted learning [24]. This highlights the demand for developing a mechanical ventilation educational program using chatbots. Therefore, this study aims to create a SDL program for mechanical ventilation nursing and evaluate its effectiveness.

Research method

Research design

This study employed a randomized controlled trial using a pretest-posttest parallel group design to develop a chatbot-based educational program for self-directed learning among nursing students and to assess its effects on clinical reasoning competency, mechanical ventilation nursing knowledge, self-confidence, and educational satisfaction.

Research participants

The participants comprised 60 fourth-year nursing students from a university in "G" province in South Korea who voluntarily participated in response to a research recruitment notice. Inclusion criteria required participants to be enrolled as fourth-year students in the Bachelor of Science in Nursing (BSN) program and to have provided voluntary consent after receiving a comprehensive explanation of the study's objectives and procedures. Exclusion criteria included prior experience with educational programs utilizing chatbot technology and possession of a registered nurse (RN) or nursing assistant certification obtained through transfer admission. The sample size calculation for assessing the effectiveness of this study was performed using the G*Power 3.1.9.2 program [25]. For a two-tail test of the difference between two independent means (two groups), with a 1:1 allocation, a power of 0.80, a significance level of 0.05, and an effect size of 0.80, the minimum required sample size for each group was determined to be 26 individuals. The choice of a relatively-large effect size of 0.8 in this study was based on previous research examining the effectiveness of chatbots targeting college students [26, 27], all of whom reported effect sizes greater than 0.8. Therefore, this study used a benchmark of 0.8, corresponding to a large effect size defined by Cohen. The participants were categorized into experimental and control groups using a 1:1 allocation through the website http://randomization.com, with block sizes of 2 and 4. Additionally, considering a dropout rate of 20.0%, 66 participants were initially selected, with 33 in each group. The flow of participants through the study is illustrated in Fig. 1.

Research tools

Clinical reasoning competency

Clinical reasoning competence was measured using a 15-item, single-factor instrument developed by Liou et al. [28] and validated in Korean by Jeong and Han [29]. This instrument was designed for the self-assessment of clinical reasoning competence among clinical nurses and graduating baccalaureate nursing students. Each item is rated on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree), with higher scores indicating a greater perceived level of clinical reasoning competence. The total score ranges from 15 to 75.

The reliability of the original instrument was Cronbach's $\alpha = 0.94$, while the Korean version showed a reliability of Cronbach's $\alpha = 0.93$. In the present study, the reliability was Cronbach's $\alpha = 0.96$.

Mechanical ventilation nursing knowledge

Mechanical ventilation-related knowledge was assessed using a 10-item tool developed by the researchers, which was validated for content validity by three nursing professors with teaching experience in respiratory and critical care nursing and two ICU nurses with over 10 years of clinical experience. Each item was scored 1 for a correct answer and 0 for an incorrect answer, with higher scores indicating a higher level of mechanical ventilation nursing knowledge. The items assessed key knowledge areas including basic principles of mechanical ventilation (2 items), classification of ventilation modes (2 items), weaning strategies and applications (2 items), complications and pathophysiological responses (2 items), alarm interpretation and situational decision-making (3 items), and endotracheal tube management and caregiver explanation (1 item). The full version of the instrument is provided as Supplemental Material (Appendix A).

Confidence in mechanical ventilation nursing

Confidence in mechanical ventilation nursing was measured using a single-item Numerical Rating Scale (NRS) ranging from 0 (not confident at all) to 10 (very confident). This simple NRS format was chosen for its practical applicability in gauging perceived self-confidence in a specific clinical skill and has been used in previous educational intervention studies.

Satisfaction in education

Satisfaction with the educational program was measured using a single-item Numerical Rating Scale (NRS) from 0 (very dissatisfied) to 10 (very satisfied). The NRS was selected to allow a straightforward evaluation of the participants' overall satisfaction with the intervention, in line with previous evaluations of nursing education programs.

Development process of the mechanical ventilation nursing chatbot educational program

Analysis phase

In the analysis phase, a search was conducted for literature related to the mechanical ventilation nursing educational program, including learning objectives and content development. This study used nine databases: PUBMED, Scopus, ProQuest, Google Scholar, and CINAHL for

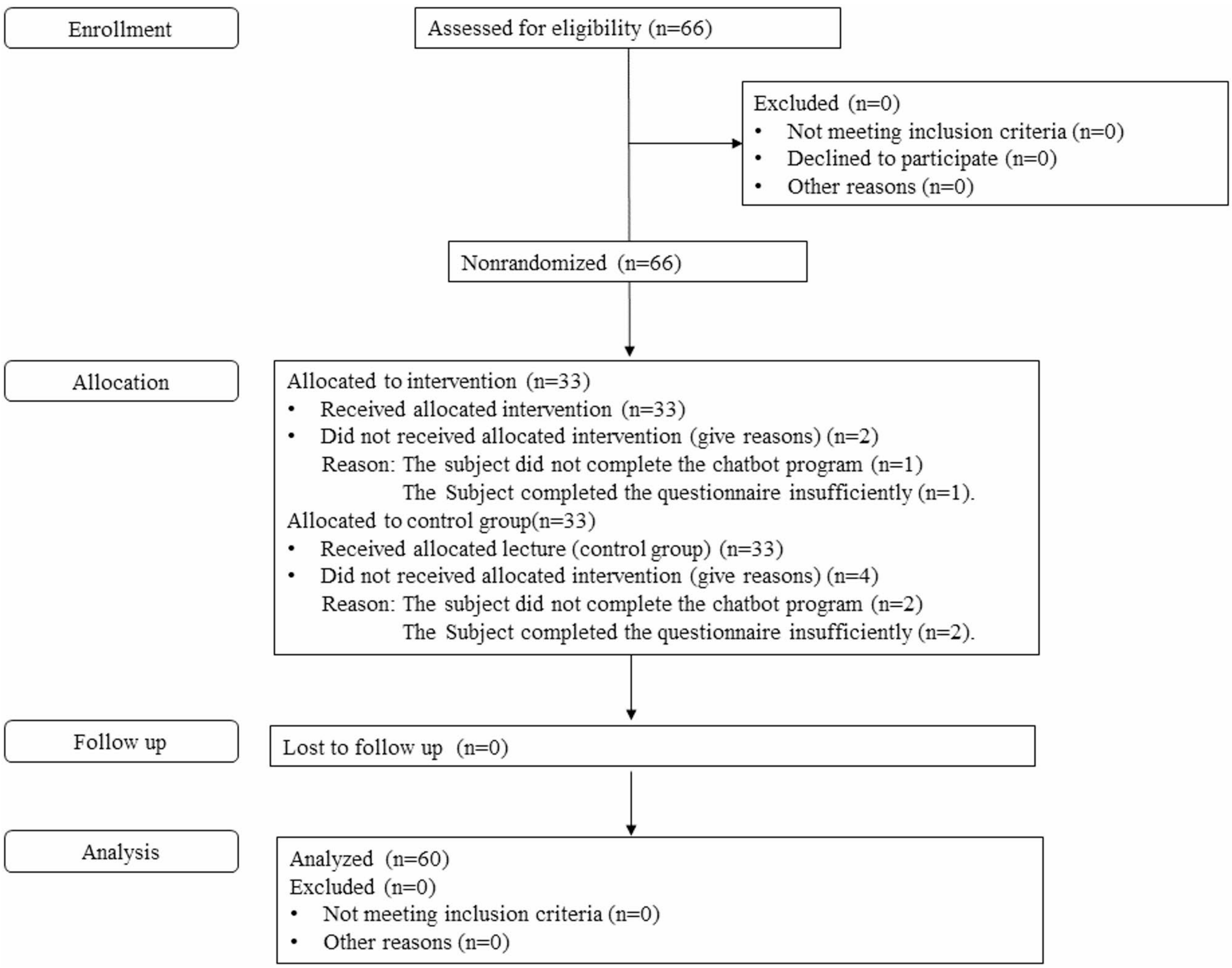

Fig. 1 Flowchart of Participant Progression Through the Study

international literature search, and DBpia, KISS, National Assembly Library, and the Korean Education and Research Information Service for domestic literature. Data extraction was initially conducted by one expert literature searcher and nursing school professor. Overall, 29 articles were extracted; after excluding 18 articles that did not contain relevant mechanical ventilation education content and one duplicate article, 10 articles were included in the final data. Based on 10 prior studies, the identified content for mechanical ventilation learning included (i) principles of mechanical ventilation, (ii) criteria for the application of mechanical ventilation, (iii) mechanical ventilation settings, (iv) modes of mechanical ventilation, (v) analysis of waveforms, (vi) verification and correction of mechanical ventilation alarms, (vii) assessment of the effectiveness of mechanical ventilation, and (viii) utilization of the Burns Wean Assessment Program for weaning from mechanical ventilation. A validity and necessity survey was conducted with two intensive care unit nurses, two nursing professors responsible for respiratory or critical care nursing, and one third-year nursing student to validate and assess the need for the learning content presented in the previous studies. Based on this feedback, an algorithm for mechanical ventilation nursing was developed.

Design phase

The chatbot is designed with components such as natural language processing (NLP), natural language understanding (NLU), and decision-making engine (DMG). NLP generates questions and answers from the chatbot, while NLU focuses on extracting meaningful information from user input. The decision-making engine is the process by which the chatbot decides whether to respond or wait [30]. This chatbot design was implemented using the LandBot online tool (https://landbot.io/), a chatbot builder. It involved configuring the program and conducting a preliminary survey with one intensive care unit nurse and nursing student each. This process made adjustments and enhancements to the appearance of the

Fig. 2 Educational Roles of Chatbots as Identified in Prior Research

screens shown to users, including the color scheme and typing speed on the chatbot interface.

Development phase

Prior to the intervention, a pre-learning video on mechanical ventilation was developed by the researchers and provided to both the experimental and control groups. The video was structured based on the learning elements identified during the analysis phase of instructional design and had a total duration of approximately 26 min. The content of the video was organized into three main sections. First, it covered ARDS and mechanical ventilation, including the definition, causes, pathophysiology, clinical symptoms, diagnostic criteria, and treatment options for ARDS. Second, it addressed mechanical ventilation modes, including the definition of mechanical ventilation, setting parameters, and the characteristics of various ventilation modes. Third, it focused on nursing care for patients receiving mechanical ventilation, covering essential nursing assessments and interventions. This video was provided equally to all participants before the intervention to ensure a consistent baseline level of knowledge for comparison between groups.

The chatbot educational program developed in this study is designed to enable SDL after receiving instruction from the instructor. This aligns with a meta-analysis conducted by Deng and Yu [31], which indicated that chatbot programs can serve three roles in education—teaching assistant, personal tutor, and learning partner (Fig. 2). This study structured the program to fulfill these three roles, facilitating SDL without requiring instructor assistance. The chatbot mechanism as a teaching assistant provides specialized knowledge on mechanical ventilation nursing and offers relevant visuals, such as photos and videos, to facilitate visual learning for students. It also provides feedback on the content. The chatbot mechanism plays the role of a learning partner engaging in text- or voice-based conversations and student interactions. Students can comfortably chat with it, like friends or family, while learning about mechanical ventilation nursing. They are encouraged by the chatbot to complete their learning and receive appreciation for correctly answering questions or progressing to the next stage, which helps them sustain their learning. As a personal tutor, the chatbot answers questions, guides students to start learning, and offers quizzes. An advantage of this chatbot educational program is that it allows students to continue their learning if they desire additional study.

The machine ventilation chatbot program consists of three phases. The structure of the chatbot interface is shown in Fig. 3. The first and second primarily focus on content for learning and acquiring knowledge. The chatbot program is designed to create a sense of familiarity with students from the beginning through initial greetings. Students also receive emoticons that move as if they are conversing with family and friends. Next, students are provided with SDL content related to machine ventilation nursing and explanations of learning objectives. Students engage in natural conversations with the chatbot, facilitating the seamless learning of content related to mechanical ventilation. The content of the machine ventilation nursing chatbot educational program is structured based on the information confirmed during the analysis phase. It is categorized into "Understanding mechanical ventilation nursing" and "Settings and modes of mechanical ventilation." Along with messages, the content includes features to review the learning material and provide students with messages of support and encouragement to enhance their SDL. Important content is emphasized to reinforce learning and retention. Furthermore, the program includes features such as photos and videos to facilitate students' learning. The learning sessions typically take 10 to 30 min, and students can engage in additional learning, which is paused if students choose not to continue. In the third stage, structured as case-based learning, students can explore and learn about nursing care for situations in which alarms occur during mechanical ventilation. There are three scenarios, each composed of high-pressure, low-pressure, and high respiration alarm cases. In these scenarios, students examine the values and alarms displayed on the mechanical ventilation screen and click to identify the reasons behind these values. They then confirm whether their answers are correct. Regardless of whether the response is correct or incorrect, they receive feedback to reinforce their learning. Additionally, for each case, students learn how to respond when an alarm goes off during mechanical ventilation.

Data collection

Data were collected from August 1–23, 2022. The researchers provided contact information to potential participants through recruitment advertisements, inviting individuals that were interested in the study to contact them directly. After the participants were recruited, the research team conducted individual meetings via the Zoom platform (Zoom Video Communications, Inc.) to provide detailed explanations of the study's objectives and the research process. The participants were categorized into the experimental and control groups using a 1:1 allocation through the website http://randomization.com, with block sizes of 2 and 4. In the experimental group. Prior to the intervention, both the experimental and control groups completed a pre-survey that assessed their general characteristics, knowledge of mechanical ventilation nursing, perceived clinical reasoning competency, self-confidence in mechanical ventilation nursing, and satisfaction with mechanical ventilation nursing education. To ensure a common baseline understanding, both groups were provided with a 26-minute pre-learning lecture video on mechanical ventilation. After viewing the video, participants entered a self-directed learning phase.

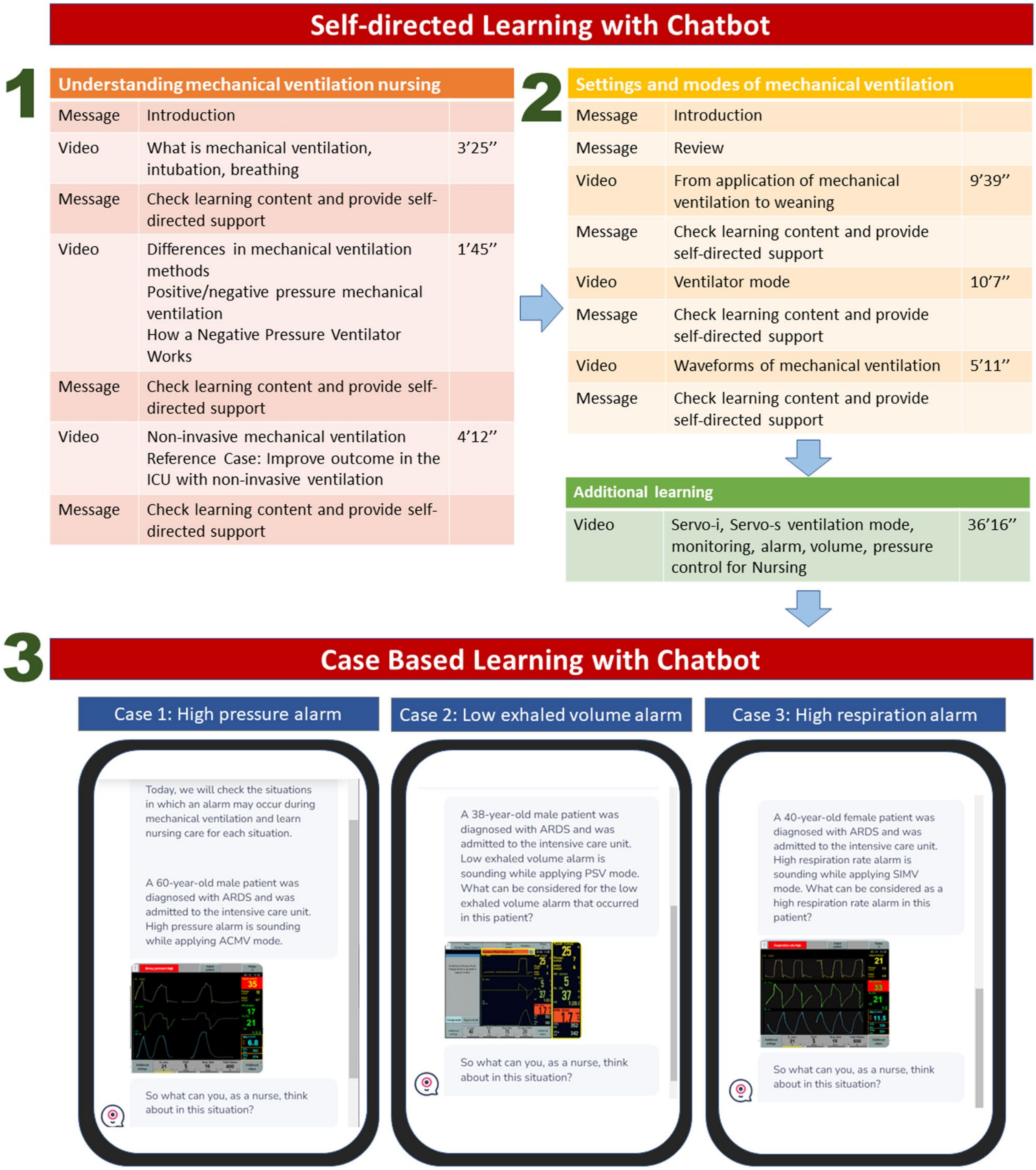

Fig. 3 Structural Design of the Machine Ventilation Chatbot Interface

During this phase, the experimental group accessed a chatbot-based educational program via a designated URL, allowing for interactive learning. In contrast, the control group was provided with relevant learning materials and URLs intended for independent study. Upon completion of the learning phase, both groups completed a post-survey using the same measurement instruments administered in the pre-survey. The post-intervention survey was conducted immediately after participants completed the educational intervention.

Data analysis

The collected data were analyzed using SPSS version 23.0 for Windows (Data Solution, Inc., Korea). The participants' general characteristics were described using frequency, percentage, mean, and standard deviation. Normality testing using the Shapiro-Wilk test was conducted for measurement variables. As appropriate, homogeneity tests for participants' characteristics and measurement variables were performed using Chi-squared tests, Fisher's exact tests, and t-tests. Independent t-tests assessed the differences between the experimental and control groups after the intervention based on the mean differences between pre- and post-intervention measurements.

Ethical considerations

All procedures in this study were conducted in accordance with the ethical standards outlined in the Dongnam Health University Institutional Review Board's (IRB) guidelines for research involving human participants. We received approval under IRB No. 1044371-202109-HR-005-01. Participants, all of whom were 18 years or older, provided informed consent before joining in the study. The consent process ensured that participants understood the nature of the study, that their participation was voluntary, and the potential risks and benefits involved. Adult participants were provided information about the study's purpose, confidentiality measures, and the right to withdraw at any time without penalty; subsequently, written consent was obtained, after which they were included in the study. The meticulous documentation of the consent process, including participants' signatures on the consent form, adhered to the approved procedures of the Dongnam Health University IRB.

In reporting this retrospective study, which involved the analysis of medical records or archived samples, we affirm that all procedures were conducted in compliance with ethical standards and guidelines specified by the Dongnam Health University's IRB or ethics committee. All data utilized in this study were obtained from fully anonymized medical records. The de-identification process was meticulously implemented to safeguard patient privacy, and no identifiable information was accessed during the analysis.

Results

Participants' general characteristics

After excluding one participant who did not complete the chatbot educational program, two who did not attend the lectures, and three with insufficient responses, a total of 60 participants were included in the final analysis—31 in the experimental group and 29 in the control group.

The participants' general characteristics are presented in Table 1. There were 17 male (28.3%) and 43 female students (71.7%). Regarding religion, 18 participants (30.0%) reported having a faith, while 42 (70.0%) had none. Clinical practice satisfaction was as follows: 34 participants were "satisfied" (56.7%), 24 were "neutral" (40.0%), and 2 were "dissatisfied" (3.3%). Satisfaction with their educational major was as follows: 38 participants were "satisfied" (63.3%), 20 were "neutral" (33.3%), and 2 were "dissatisfied" (3.3%). Satisfaction with university life was as follows: 32 participants were "satisfied" (53.3%), 26 were "neutral" participants (43.3%), and 2 "dissatisfied" (3.3%). The analysis of homogeneity between groups regarding the general characteristics of the study participants showed no statistically significant differences. The results of the Shapiro–Wilk test indicated that all continuous variables were normally distributed (p >.05).

Table 1 Homogeneity test of general characteristics of subjects (N = 61)

Variables	Category	Experimental (n = 31) *n* (%)	Control (n = 29) *n* (%)	Total (N = 60) *n* (%)	χ^2/t	*p*-value
Age (year)		22.68 ± 1.33	23.28 ± 1.13	22.97 ± 1.26	-1.88	0.066
Gender	Male	7 (11.7)	10 (16.7)	17 (28.3)	1.05	0.394[†]
	Female	24 (40.0)	19 (31.7)	43 (71.7)		
Religion	Yes	10 (16.7)	8 (13.3)	18 (30.0)	0.16	0.693
	No	21 (35.0)	21 (35.0)	42 (70.0)		
Satisfaction with their Clinical practice	Satisfied	21 (35.0)	13 (21.7)	34 (56.7)	3.32	0.190
	Somewhat satisfied	9 (15.0)	15 (25.0)	24 (40.0)		
	Not satisfied	1 (1.7)	1 (1.7)	2 (3.3)		
Satisfaction with their major	Satisfied	21 (35.0)	17 (29.3)	38 (63.3)	0.56	0.788
	Somewhat satisfied	9 (15.0)	11 (18.3)	20 (33.3)		
	Not satisfied	1 (1.7)	1 (1.7)	2 (3.3)		
Satisfaction with their university life	Interested	18 (30.0)	14 (23.3)	32 (53.3)	0.59	0.796
	Somewhat interested	12 (20.0)	14 (23.3)	26 (43.3)		
	No interested	1 (1.7)	1 (1.7)	2 (3.3)		

[†]: Fisher's exact test

Table 2 Homogeneity test of dependent variables ($N=61$)

Variables	Group	M ± SD	t	*p*-value
Knowledge	Experimental	6.97 ± 1.64	1.81	0.075
	Control	6.41 ± 1.50		
Clinical reasoning competency	Experimental	43.68 ± 11.63	-1.03	0.308
	Control	46.55 ± 9.89		
Confidence	Experimental	4.35 ± 1.99	0.38	0.706
	Control	4.17 ± 1.17		
Satisfaction in education	Experimental	5.65 ± 2.23	-0.64	0.525
	Control	6.00 ± 2.05		

Table 3 Effects on education program ($N=61$)

Variables	Group	Pre-test (M ± SD)	Post-test (M ± SD)	t	*p*
Clinical reasoning competency	Exp	43.68 ± 11.63	56.10 ± 9.26	-5.00	< 0.001
	Con	46.55 ± 9.89	47.21 ± 5.51		
Knowledge	Exp	6.97 ± 1.64	7.32 ± 1.22	-0.09	0.926
	Con	6.41 ± 1.50	6.72 ± 1.33		
Confidence	Exp	4.35 ± 1.99	6.13 ± 1.67	-2.62	0.011
	Con	4.17 ± 1.71	4.76 ± 1.35		
Satisfaction in education	Exp	5.65 ± 2.23	7.42 ± 1.59	-3.51	< 0.001
	Con	6.00 ± 2.05	5.55 ± 1.74		

M = Mean; SD = Standard deviation; Exp = Experimental Group; Con = Control Group

Homogeneity testing

The results of homogeneity testing between the experimental and control group for the measured variables showed no statistically significant differences in Clinical reasoning competency (p =.308), Knowledge (p =.075), Confidence (p =.706), and Satisfaction in education (p =.525) (Table 2).

Statistical hypothesis testing

The results of the difference tests for the measurement variables between the experimental and control groups are as follows (Table 3). The experimental group, which received the chatbot educational program, showed statistically significant higher levels of Clinical reasoning competency (t = -5.00, p <.001), Confidence (t = -2.62, p =.011), and Satisfaction in education (t = -3.51, p <.001) compared to the control group. However, the two groups had no statistically significant difference in Knowledge (t = -0.09, p =.926).

Discussion

This study aims to develop a chatbot educational program incorporating SDL for nursing students to learn mechanical ventilation nursing and to assess its effectiveness. The study focuses on the following aspects.

First, clinical reasoning competency differed significantly between the experimental and control groups. This contrasts with a study on a chatbot educational program for fetal health assessment, which found no significant difference in clinical reasoning competency between the two groups [16]. These findings are consistent with previous studies in which a VR-based mechanical ventilation training program was shown to improve nursing students' clinical reasoning competency [32]. Similarly, a study by Salameh et al. [5] reported that high-fidelity simulation (HFS)–based mechanical ventilation education significantly improved clinical decision-making abilities among nursing students. This difference may be attributed to the fact that the chatbot educational program developed in this study incorporated case-based learning methods, allowing students to utilize clinical judgment effectively. In the context of nursing practice, clinical reasoning involves the cognitive process of problem-solving for individual patient care in various stages of disease prevention, diagnosis, or treatment. Rational decision-making requires consideration of multiple aspects of patient care. The chatbot educational program developed in this study may have been helpful for nursing students to critically question and seek answers while caring for patients in situations with various mechanical ventilation monitoring alarms. Through the stages of the chatbot educational program, students acquired knowledge and interpreted data based on this knowledge. They made clinical judgments on what nursing actions to take by considering patient symptoms and test results. This process likely enhanced their clinical reasoning competency.

Second, in this study, mechanical ventilation nursing knowledge increased in both the experimental and control groups from pre-test to post-test; however, no statistically significant difference was observed between the groups. This result is consistent with a previous study by Han, Park, and Lee [16], in which electronic fetal monitoring education for nursing students led to knowledge improvement in both groups without a significant difference between them. Meanwhile, Huang et al. [24] reported that chatbot-based education had a significant positive effect on learners' knowledge acquisition. Although the present study also employed a chatbot structure that included repeated interactions, formative quizzes, and personalized feedback, no significant difference in knowledge improvement was found between the experimental and control groups. This may be because the pre-learning video provided to both groups sufficiently strengthened their foundational theoretical knowledge, potentially equalizing the learning levels across groups. Additionally, the knowledge assessment tool used in this study was based on simple correct-or-incorrect responses, which may not have been sensitive enough to detect subtle differences in understanding depending on the learning method. Therefore, future research should consider incorporating tools that assess higher-order thinking and tracking the long-term impact of chatbot learning in terms of repetition and depth of engagement.

Third, there was a significant difference in self-confidence in mechanical ventilation nursing between the experimental and control groups. This result is consistent with a Taiwanese study, which applied chatbots to nursing students and found that it could enhance self-efficacy [22]. In that study, students generally reported that learning through mobile chatbots enhanced their engagement, performance, and self-efficacy. Furthermore, similar to a study that applied a chatbot program for history-taking assessment to nursing students [33], it showed that the chatbot program allowed students to practice repetitively, providing them with more opportunities to attempt and learn from mistakes, ultimately boosting their confidence. For introverted students who struggle to reinforce their theoretical knowledge, chatbots could reduce the stress associated with trial and error. It also enhances their self-confidence, as indicated by the research findings. Therefore, utilizing chatbot programs can help boost the self-confidence of nursing students.

Fourth, satisfaction with mechanical ventilation nursing education showed a significant difference between the experimental and control groups. This aligns with prior research [34], which examined the impact of chatbot integration with AI on students' online learning experiences, indicating a positive effect. Additionally, it is consistent with studies [24, 35], which found that chatbot education positively impacted learners' satisfaction with the learning experience. Chatbots allow learners to engage in conversations anytime and anywhere, free from time and space constraints. Learners can interact with chatbots, lower tension and stress levels, and engage in repetitive learning [36]. This flexibility and reduced pressure contributed to the improvement in satisfaction. In a Taiwanese study, where chatbots were applied to nursing students' learning [22], users reported that they could concentrate on their lessons, found chatbot learning enjoyable, and were able to explore relevant information actively. The chatbot educational program in this study also leveraged the advantages of various chatbots, enabling SDL, which is believed to have contributed to these results.

This study has several limitations. First, the same knowledge assessment tool was used for both pre- and post-intervention measurements, which may have introduced a memory effect. However, using the same instrument also ensured consistency and reliability across time points. Second, confidence and satisfaction were measured using a single-item 0–10 Numerical Rating Scale (NRS). Although the NRS is widely used in educational and clinical settings for its simplicity, it is not a validated multi-item scale, which may limit the depth and psychometric precision of the measurement. Third, although both groups were provided with the same pre-learning video, the actual time participants spent engaging with the educational materials during the self-directed learning phase was not measured, which may have influenced the outcomes.

Future studies should consider using validated measurement instruments and quantifying learning engagement to enable more accurate evaluation of educational effects.

Conclusions and suggestions

This study found that the chatbot-assisted mechanical ventilation nursing education program enhanced nursing students' clinical reasoning competency, confidence, and satisfaction. Applying chatbot-based education to critical care nursing education for students and novice nurses may lead to improved clinical judgment in situations related to mechanical ventilation, fostering confident healthcare professionals. This will ultimately contribute to establishing trust and respect between nurses and patients in an intensive care unit, ensuring patient safety, and promoting optimal health outcomes in a hospital setting. However, it should be noted that this study evaluated the effectiveness with a single university as the sample, and there is a limitation in that the post-assessment of measurement variables was conducted only once. For future research, it is recommended to employ a longitudinal study design to assess the medium to long-term effects and explore the in-depth impact of the chatbot educational program through qualitative research analysis.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12909-025-07316-2.

Supplementary Material 1

Acknowledgements

Not applicable.

Author contributions

Conceptualization or/and Methodology: Han JW, Lee H, Park J; Data curation or/and Analysis: Han JW, Lee H, Park J; Investigation: Han JW, Lee H, Park J; Resources or/and Software: Han JW, Lee H, Park J; Validation: Han JW, Lee H, Park J; Visualization: Han JW, Lee H, Park J; Writing: original draft or/and review & editing: Han JW, Lee H, Park J. All authors read and approved the final manuscript.

Funding

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (grant number: 2022R1I1A3072862).

Data availability

The data that support the findings of this study are available on request from the corresponding author.

Declarations

Ethics approval or consent to participate

All procedures in this prospective study involving human participants were conducted in accordance with the ethical standards outlined in the Dongnam Health University Institutional Review Board's (IRB) guidelines. The study received approval under IRB No. 1044371-202109-HR-005-01. All participants, who were 18 years or older, provided informed consent before participating in the study. The informed consent process ensured that participants fully understood the nature of the study, that their participation was voluntary, and the potential risks and benefits involved. Participants were informed about the study's purpose, confidentiality measures, and their right to withdraw from the study at any time without penalty. Written consent was obtained before their inclusion in the study, and the documentation, including signatures, adhered to the procedures approved by the Dongnam Health University IRB. For the retrospective component of this study, which involved the analysis of medical records or archived samples, we confirm that all procedures were conducted in compliance with the ethical standards and guidelines set by the Dongnam Health University's IRB. All data were obtained from fully anonymized medical records, and the de-identification process was rigorously implemented to protect patient privacy. No identifiable information was accessed during the data analysis. In addition to the institutional ethical standards, all procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation and with the Helsinki Declaration of 1975, as revised in 2000. Informed written consent to participate was obtained from all students, and participation was voluntary. The study was conducted in accordance with the principles of the Declaration of Helsinki.

Consent for publication

Written informed consent was obtained from all participants to publish the data.

Competing interests

The authors declare no competing interests.

Received: 13 October 2024 / Accepted: 8 May 2025
Published online: 02 June 2025

References

1. Lasater K. Clinical judgment: the last frontier for evaluation. Nurse Educ Pract. 2011;11:86–92.
2. Mahase E. Covid-19: most patients require mechanical ventilation in first 24 hours of critical care. BMJ. 2020;368:m1201.
3. Williams G, Schmollgruber S, Alberto L. Consensus forum: worldwide guidelines on the critical care nursing workforce and education standards. Crit Care Clin. 2006;22:393–406.
4. Lee YB, Kim YJ. Clinical nurses' awareness and learning needs of education of core basic nursing skills. Asia Pac J Multimed Serv Converg Art Humanit Sociol. 2019;9:705–16.
5. Salameh B, Ayed A, Kassabry M, Lasater K. Effects of a complex case study and high-fidelity simulation on mechanical ventilation on knowledge and clinical judgment of undergraduate nursing students. Nurse Educ. 2021;46:E64–9.
6. Turrise SL, Thompson CE, Hepler M. Virtual simulation: comparing critical thinking and satisfaction in RN-BSN students. Clin Simul Nurs. 2020;46:57–61.
7. Corbridge SJ, Robinson FP, Tiffen J, Corbridge TC. Online learning versus simulation for teaching principles of mechanical ventilation to nurse practitioner students. Int J Nurs Educ Scholarsh. 2010;7:Article12.
8. Doolen J, Mariani B, Atz T, Horsley TL, Rourke JO, McAfee K, et al. High-fidelity simulation in undergraduate nursing education: a review of simulation reviews. Clin Simul Nurs. 2016;12:290–302.
9. Foronda CL, Fernandez-Burgos M, Nadeau C, Kelley CN, Henry MN. Virtual simulation in nursing education: a systematic review spanning 1996–2018. Simul Healthc. 2020;15:46–54.
10. Dehghan-Nayeri N, Vasli P, Seylani K, Fallahi S, Rahimaghaee F, Kazemnejad A. The effectiveness of workshop and multimedia training methods on the nurses' decision-making skills regarding weaning from mechanical ventilation. Dimens Crit Care Nurs. 2020;39:91–100.
11. Baber H. Determinants of students' perceived learning outcome and satisfaction in online learning during the pandemic of COVID19. J Educ Elearn Res. 2020;7:285–92.
12. Loeng S. Self-directed learning: a core concept in adult education. Educ Res Int. 2020;2020:1–12.
13. Kim D, Jung E, Yoon M, Chang Y, Park S, Kim D, et al. Exploring the structural relationships between course design factors, learner commitment, self-directed learning, and intentions for further learning in a self-paced MOOC. Comput Educ. 2021;166:104171.
14. Lee MK, Park H. Exploring factors influencing usage intention of chatbot–chatbot in financial service. J Korean Soc Qual Manag. 2019;47:755–65.
15. Chen HL, Vicki Widarso G, Sutrisno H. A chatbot for learning Chinese: learning achievement and technology acceptance. J Educ Comput Res. 2020;58:1161–89.
16. Han JW, Park J, Lee H. Analysis of the effect of an artificial intelligence chatbot educational program on non-face-to-face classes: a quasi-experimental study. BMC Med Educ. 2022;22:830.
17. Kuhail MA, Thomas J, Alramlawi S, Shah SJH, Thornquist E. Interacting with a chatbot-based advising system: Understanding the effect of chatbot personality and user gender on behavior. Informatics. 2022;9:81.
18. Kuosmanen L, Vartiainen A-K, Nieminen H, Kostenius C, Bond R, Mulvenna M, et al. Development process of artificial intelligence based chatbot to support and promote mental wellbeing in sparsely populated areas of five European countries. Eur Psychiatr. 2022;65:S168.
19. Hungerbuehler I, Daley K, Cavanagh K, Garcia Claro H, Kapps M. Chatbot-based assessment of employees' mental health: design process and pilot implementation. JMIR Form Res. 2021;5:e21678.
20. Zhang J, Oh YJ, Lange P, Yu Z, Fukuoka Y. Artificial intelligence chatbot behavior change model for designing artificial intelligence chatbots to promote physical activity and a healthy diet: viewpoint. J Med Internet Res. 2020;22:e22845.
21. Piao M, Kim J, Ryu H, Lee H. Development and usability evaluation of a healthy lifestyle coaching chatbot using a habit formation model. Healthc Inf Res. 2020;26:255–64.
22. Chang CY, Hwang GJ, Gau ML. Promoting students' learning achievement and self-efficacy: a mobile chatbot approach for nursing training. Br J Educ Technol. 2022;53:171–88.
23. Buchanan C, Howitt ML, Wilson R, Booth RG, Risling T, Bamford M. Predicted influences of artificial intelligence on the domains of nursing: scoping review. JMIR Nurs. 2020;3:e23939.
24. Huang W, Hew KF, Gonda DE. Designing and evaluating three chatbot-enhanced activities for a flipped graduate course. Int J Mech Eng Robot Res. 2019;8:813–8.
25. Faul F, Erdfelder E, Lang AG, Buchner AG. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods. 2007;39:175–91.

26. Abbasi S, Kazi H. Measuring effectiveness of learning chatbot systems on student's learning outcome and memory retention. Asian J Appl Sci Eng. 2014;3:251–60.
27. Fryer LK, Thompson A, Nakao K, Howarth M, Gallacher A. Supporting self-efficacy beliefs and interest as educational inputs and outcomes: framing AI and human partnered task experiences. Learn Individ Differ. 2020;80.
28. Liou SR, Liu HC, Tsai HM, Tsai YH, Lin YC, Chang CH, et al. The development and psychometric testing of a theory-based instrument to evaluate nurses' perception of clinical reasoning competence. J Adv Nurs. 2016;72:707–17.
29. Joung J, Han JW. Validity and reliability of a Korean version of nurse clinical reasoning competence scale. J Korea Acad Ind Coop Soc. 2017;18:304–10.
30. Adamopoulou E, Moussiades L. An overview of chatbot technology. In: Maglogiannis I, Iliadis L, Pimenidis E, editors. Artificial intelligence applications and innovations. AIAI 2020. IFIP Advances in Information and Communication Technology. Volume 584. Cham: Springer; 2020. pp. 373–83.
31. Deng X, Yu Z. A meta-analysis and systematic review of the effect of chatbot technology use in sustainable education. Sustainability. 2023;15:2940.
32. Lee H, Han JW. Development and evaluation of a virtual reality mechanical ventilation education program for nursing students. BMC Med Educ. 2022;22:775.
33. Chen Y, Lin Q, Chen X, Liu T, Ke Q, Yang Q, et al. Need assessment for history-taking instruction program using chatbot for nursing students: a qualitative study using focus group interviews. Digit Health. 2023;9:20552076231185435.
34. Deveci Topal A, Dilek Eren C, Kolburan Geçer A. Chatbot application in a 5th grade science course. Educ Inf Technol (Dordr). 2021;26:6241–65.
35. Chang CY, Kuo SY, Hwang GH. Chatbot-facilitated nursing education: incorporating a knowledge-based chatbot system into a nursing training program. Educ Technol Soc. 2022;25:15–27.
36. Kuhail MA, Alturki N, Alramlawi S, Alhejori K. Interacting with educational chatbots: a systematic review. Educ Inf Technol. 2023;28:973–1018.

Publisher's note