

Assessment of Student Learning Outcomes Using AI and Simulation in Choking Management Skills: A Pilot Randomized Controlled Trial

Janice C. Palaganas^{1, 2, ✉}, Alex Morton³, Fatima Adem⁴, Jianna Ramos⁵, Arjun Kumar⁶, Khaulah Jawed⁷, Maria Bajwa^{8, 1, 2, 9}

1. *REBEL Lab, Center of Excellence in Healthcare Simulation Research, MGH Institute of Health Professions, Boston, USA*
2. *Health Professions Education, MGH Institute of Health Professions, Boston, USA*
3. *Center of Excellence in Healthcare Simulation Research, MGH Institute of Health Professions, Boston, USA*
4. *Emergency Medicine, University of Michigan, Ann Arbor, USA*
5. *Biochemistry, Case Western Reserve University, Cleveland, USA*
6. *Biology and Community Health, Tufts University, Medford, USA*
7. *Health Professions Education, NUST School of Health Sciences, Islamabad, PAK*
8. *College of Health Professions, Upstate Medical University, Syracuse, USA*
9. *Simulation, Nova Southeastern University, Fort Lauderdale, USA*

Received: March 23, 2026 | Review began: March 26, 2026 | Review ended: May 28, 2026 | Published: June 02, 2026

© Copyright 2026

This is an open access article distributed under the terms of the Creative Commons Attribution License CC-BY 4.0., which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Background

As the integration of artificial intelligence (AI) into health professions education becomes increasingly inevitable, educators face the challenge of determining how to use AI tools effectively and safely. Despite the potential of AI-assisted self-directed learning, limited evidence exists comparing its impact to traditional methods, especially regarding psychomotor and critical thinking skills in emergency care training.

Methods

This pilot randomized controlled trial compared AI-assisted self-directed learning, lecture-based instruction, and simulation-based education in teaching emergency choking management to pre-professional health students. Twenty students ($n = 6-7$ per group) enrolled in a two-week preparatory program were randomly assigned to one of three instructional groups: lecture, simulation, or AI-assisted self-directed learning using ChatGPT-4o. Notably, the lecture and simulation groups received roughly 60-minute faculty-led sessions, while the AI-assisted group received the standardized educational outline and was asked to self-study using ChatGPT-4o. All participants completed a post-intervention simulation scenario and a knowledge assessment on the third day of the two-week program (multiple-choice and essay), followed by a repeat assessment on the final day of the program (nine days later). Simulation performance was scored using an expert-developed checklist. Essay responses and AI interaction transcripts were evaluated using a modified Holistic Critical Thinking Scoring Rubric. A focus group was conducted to explore learner experiences.

Results

Twenty pre-professional students participated (lecture $n = 7$, simulation $n = 6$, AI-assisted $n = 7$). Students in the lecture and simulation groups demonstrated higher performance and confidence levels than those in the AI-assisted group, with statistically significant differences on performance and observer-rated confidence (inter-rater reliability $\alpha = .83$). Analysis of AI transcripts revealed substantial variability in prompting strategies, metacognitive engagement, and depth of

How to cite this article:

Palaganas J C, Morton A, Adem F, et al. (June 02, 2026) Assessment of Student Learning Outcomes Using AI and Simulation in Choking Management Skills: A Pilot Randomized Controlled Trial. *Cureus J Comput Sci* 3 : es44389-026-00094-y. DOI <https://doi.org/10.7759/s44389-026-00094-y>

learning. Focus group data indicated both perceived benefits and psychological barriers to AI use, including feelings of uncertainty and guilt.

Conclusion

In this small pilot, AI-assisted self-directed learning appeared to underperform faculty-supported lecture and simulation formats in teaching emergency care skills. Because the AI group received less instructional support than the lecture and simulation groups, these findings should be interpreted as a comparison of AI as a self-directed modality against faculty-led instruction, rather than AI-assisted teaching. The data are suggestive that early development of AI literacy, prompting competency, and psychological safety may be important for successful integration. Future research should involve larger, more diverse cohorts, baseline assessment, comparable instructional support across arms, and explicit study of AI literacy and prompt-design scaffolding for blended instructional models.

Categories: AI applications, Evaluation Methods, Safe and trustworthy AI in healthcare

Keywords: artificial intelligence (ai), health professions education, simulation-based education, critical thinking, choking

Introduction

Global exposure to large language models (LLMs) began in 2021 and rapidly expanded with the public release of tools like ChatGPT in late 2022, transforming classrooms, projects, and daily communication [1]. By early 2023, generative AI was reshaping healthcare, education, and social domains [2]. While LLMs now assist with clinical reasoning and decision support [3] and have promising potential as virtual learning assistants for self-directed learning [4], they also raise issues of academic integrity, data privacy, and ethics [5,6]. Institutions worldwide are calling for modernization of teaching and assessment methods to address these challenges [2]. Although AI is transforming educational practice, its implications for learning outcomes and assessment remain underexplored [7,8].

Competency-based education shifts focus from content delivery to performance outcomes assessed in authentic contexts. Healthcare simulation, for instance, has proven more effective than traditional exams in demonstrating knowledge retention, critical thinking, and clinical competence [9]. As generative AI becomes ubiquitous, it adds new complexity: many students enter higher education already accustomed to AI use, shaping their learning strategies and beliefs about knowledge.

High school or “pre-professional” students, poised to enter health professions programs, are a critical population for early investigation. Their reliance on conversational AI to generate, refine, and explore ideas presents a novel opportunity to study thinking processes, including critical thinking, within AI-mediated learning environments [8,10,11]. Understanding how these learners engage with AI and how their interactions reflect reasoning and judgment can inform the development of higher education interventions.

Given that large-scale AI integration into health professions curricula is imminent, early studies at the pre-professional level can guide responsible design of higher education models. Pre-professional students’ technological proficiency also makes them ideal for exploring effective and ethical uses of AI in clinical education.

This study sits at the intersection of AI and simulation as tools for both learning and assessment. We aim to evaluate how AI-assisted self-directed learning compares to lecture and simulation methods in developing clinical performance, knowledge retention, and critical thinking, which we define for the purpose of this study as “the process of making purposeful, reflective and fair-minded judgments about what to believe or what to do” [12,13]. We deliberately compared AI-assisted self-directed learning against both lecture and simulation as these three modalities represent the most common observed formats encountered by pre-professional learners as they enter health professions programs: didactic transfer of content (lecture), experiential practice with feedback (simulation), and increasingly, AI-mediated self-study (reference). While lecture and simulation are pedagogically distinct, both share the feature of faculty-led design and oversight, and pre-professional students routinely encounter both during their early training; we therefore included both

How to cite this article:

Palaganas J C, Morton A, Adem F, et al. (June 02, 2026) Assessment of Student Learning Outcomes Using AI and Simulation in Choking Management Skills: A Pilot Randomized Controlled Trial. *Cureus J Comput Sci* 3 : es44389-026-00094-y. DOI <https://doi.org/10.7759/s44389-026-00094-y>

as comparators to position AI alongside the realistic mix of formats students experience. Additionally, this study examines the feasibility of combining AI-generated transcripts (e.g., chat or verbal interactions) with traditional exams to assess knowledge and reasoning, and explores the advantages and challenges of integrating simulation performance, knowledge tests, and AI chat threads in assessing learning outcomes. This work represents an initial step toward a research agenda assessing knowledge, reasoning, and competency through AI-integrated simulation.

Materials And Methods

Methods

This pilot randomized controlled trial used a parallel three-arm design [14] involving 20 pre-professional students voluntarily enrolled in a two-week intensive summer program for aspiring healthcare providers at a small university in Boston, Massachusetts.

This study was deemed exempt by the Mass General Brigham (MGB) Institutional Review Board (Protocol ID 2025P001575).

Participants were randomly assigned via computer-generated allocation to one of three instructional groups: lecture-based, simulation-based, or AI-assisted self-learning. Participation was voluntary and was not tied to course grading or academic standing in the summer program; students were informed verbally and in writing prior to consent that (a) declining to participate would have no impact on their participation in the broader summer program, (b) data would be reported in aggregate and de-identified, and (c) they could withdraw at any point without consequence. The lecture and simulation arms each consisted of a single approximately 60-minute faculty-led session delivered by the same core teaching team to ensure content consistency. The lecture session was a didactic presentation that combined PowerPoint slides with visualization of choking-management maneuvers; it was delivered by an experienced emergency physician faculty member, and it did not include hands-on student practice. The simulation session followed the same standardized educational outline and included pre-briefing [15], scenario-based practice with multiple low-fidelity manikins (adult, pregnant adult, child, and infant), and a plus-delta debriefing led by a simulation-trained faculty member. The AI-assisted self-learning group was provided the same standardized educational outline (Appendix I) and individual paid ChatGPT-4o accounts [16], and was instructed to use ChatGPT-4o to learn the content during a session of comparable scheduled duration. A faculty member was in the room to provide technical and session support. Importantly, no formal scaffolding in prompt design or AI literacy was provided; students were free to ask ChatGPT for explanations, examples, photos, or videos as they saw fit. This decision was intentional, in order to characterize how pre-professional students engage with a generative AI tool when given access without prior instruction in prompting; we discuss the implications of this design decision for the comparability of arms in our Limitations. All participants in the AI group submitted their full ChatGPT-4o transcripts for analysis. Faculty collaborated to ensure content consistency across all instructional modalities.

Following instruction, all students completed an individual infant choking simulation and a knowledge assessment consisting of multiple-choice, open-ended, and essay questions (Appendix III). During this assessment, students in all three arms were permitted to use ChatGPT-4o mini in anonymous mode (so that responses would not be personalized to prior chat history) so that we could examine how learners across modalities engaged with AI when answering questions. All participants provided both their test responses and AI chat records. An exploratory focus group was conducted after testing, followed by a repeat post-test without AI access nine days later (at the end of the two-week program) to measure retention and independent knowledge performance.

How to cite this article:

Palaganas J C, Morton A, Adem F, et al. (June 02, 2026) Assessment of Student Learning Outcomes Using AI and Simulation in Choking Management Skills: A Pilot Randomized Controlled Trial. *Cureus J Comput Sci* 3 : es44389-026-00094-y. DOI <https://doi.org/10.7759/s44389-026-00094-y>

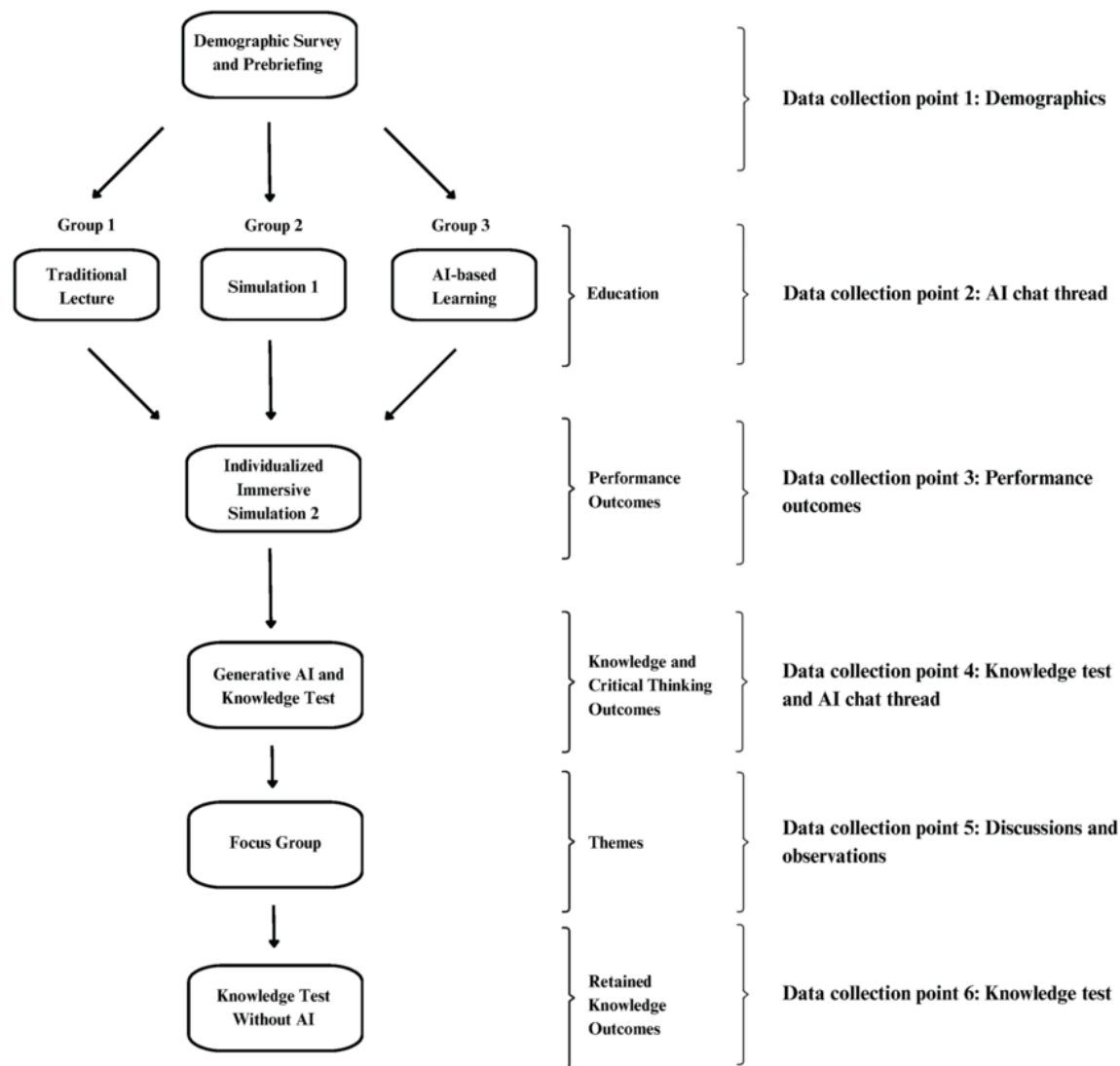


FIGURE 1: Study Schema

Instruments

The Choking Simulation Case involved a standardized mother and her 8-month-old manikin infant [15]. Because choking management guidelines differ for adults but generally have the same guidelines for infants less than 1 year old, we chose an infant scenario (Appendix I). While taking the mother's medical history, the infant began choking, and the student was expected to perform choking management. A recorded baby cry signaled recovery and scenario completion. There was no debriefing as we sought only to measure performance skills.

The Choking Simulation Performance Measure Tool was created using course content and American Heart Association guidelines [17,18] of choking management (Appendix II).

How to cite this article:

Palaganas J C, Morton A, Adem F, et al. (June 02, 2026) Assessment of Student Learning Outcomes Using AI and Simulation in Choking Management Skills: A Pilot Randomized Controlled Trial. *Cureus J Comput Sci* 3 : es44389-026-00094-y. DOI <https://doi.org/10.7759/s44389-026-00094-y>

The Choking Knowledge Test was developed by content experts and was based on the course content with essay questions that intentionally included new content to provide space for ChatGPT searches and critical thinking (Appendix III). This test was the same in Data Collection Points 4 and 6 (Figure 7).

The Holistic Critical Thinking Scoring Rubric (HCTSR) evaluates overall critical thinking quality using a four-point scale based on clarity, relevance, depth, logic, and fairness [13]. Originally designed for written or verbal responses [13,19,20], it was adapted in this study with an added column to assess critical thinking demonstrated in AI chat transcripts (Appendix IV).

One semi-structured focus group that included all students was conducted using two primary topics: 1) How was your experience using ChatGPT with your exam? Why? 2) Did it help or not help you? How?

Analysis

Demographic data: Descriptive statistics were calculated for demographic and outcome variables. A sentiment analysis was performed on the open-ended demographic survey questions using Dedoose (Version 10.0.35) [21] to understand whether students generally liked, disliked, or were neutral about each method of teaching (see Table 2 for results).

Performance data: Following interrater reliability training using two videos, performance measures were scored independently by four researchers from video recordings: two researchers were observers; two were the standardized patients.

Knowledge test and AI chat threads data: Multiple choice answers were analyzed descriptively. Following interrater reliability training using two tests, essay questions in combination with the AI Chat threads were scored independently by two researchers using the HCTSR via consensus. Pearson correlation and multiple regression were then used to examine associations in outcomes.

Focus group data: Qualitative content analysis was conducted using Dedoose on the focus group interview transcript by two researchers via consensus [22].

Follow-up knowledge test data: The essays in the follow-up test were scored by the same two researchers who scored the initial knowledge test. There were no AI Chat threads for the follow-up test.

Results And Discussion

Results

A total of 20 pre-professional students were enrolled and randomly assigned to one of three groups (lecture $n = 7$, simulation $n = 6$, AI-assisted $n = 7$). Participants across groups were similar in grade and prior AI usage. The AI group, by chance of randomization, was slightly older, entirely female, more racially diverse, and had the highest proportion of English as an additional language (EAL) students; we discuss the implications of this distribution for interpretation in the Limitations. Table 1 describes the participant demographics.

How to cite this article:

Variable	Lecture (n = 7)	Simulation (n = 6)	AI (n = 7)
Age (M ± SD, Range)	15.3 ± 0.5, 15–16	16.0 ± 0.0, 16–16	16.3 ± 0.8, 15–17
Grade (M ± SD, Range)	10.6 ± 1.1, 10–12	11.4 ± 0.5, 11–12	11.2 ± 0.8, 10–12
Self-identified Gender			
Female (%)	3 (43%)	4 (67%)	7 (100%)
Male (%)	3 (43%)	2 (33%)	0 (0%)
Non-binary (%)	1 (14%)	0 (0%)	0 (0%)
Self-identified Race/Ethnicity			
Asian (%)	4 (57%)	2 (33%)	2 (29%)
White (%)	3 (43%)	3 (50%)	5 (71%)
Hispanic (%)	1* (14%)	2 (33%)	3 (43%)
Black (%)	0 (0%)	1 (17%)	1 (14%)
English as primary or additional language			
English as primary language, Yes (%)	6 (86%)	6 (100%)	5 (71%)
English as primary language, No (%)	1 (14%)	0 (0%)	2 (29%)
AI Use			
Has used AI, Yes (%)	7 (100%)	6 (100%)	7 (100%)
Has used AI, No (%)	0 (0%)	0 (0%)	0 (0%)

TABLE 1: Demographics by Groups

*Multi-racial cases.

Simulation consistently outperformed the other groups on the performance checklist, followed closely by Lecture, with the AI group scoring lower overall; pairwise comparisons reached statistical significance for AI versus the two faculty-led arms (see Figure 2 and Table 2). Lecture and Simulation showed slight, non-statistically significant declines in knowledge from post-intervention to follow-up, with Simulation remaining the highest overall. The AI group demonstrated a delayed

How to cite this article:

Palaganas J C, Morton A, Adem F, et al. (June 02, 2026) Assessment of Student Learning Outcomes Using AI and Simulation in Choking Management Skills: A Pilot Randomized Controlled Trial. Cureus J Comput Sci 3 : es44389-026-00094-y. DOI <https://doi.org/10.7759/s44389-026-00094-y>

improvement in critical thinking from post-intervention to follow-up (mean 1.57 to 3.00), nearing Lecture levels but with high variability (SD = 2.38). This pattern is suggestive of a large effect size, reflecting substantial average improvement, but with inconsistent impact across participants.

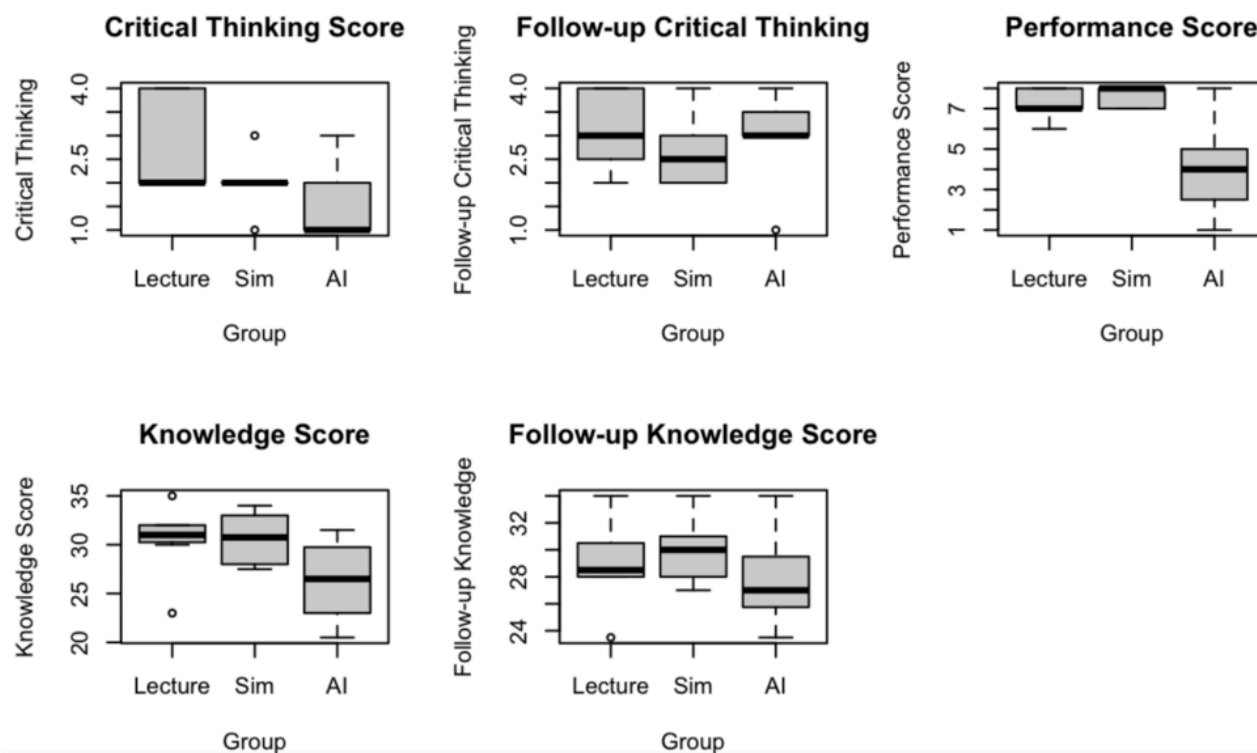


FIGURE 2: Outcomes by Groups

Boxplots illustrating group differences for the five primary study outcomes: Critical Thinking Score, Follow-up Critical Thinking Score, Knowledge Score, Follow-up Knowledge Score, and Performance Score. Each plot compares the distribution of scores among the Lecture, Simulation, and AI-assisted groups. Boxes represent the interquartile range (IQR); horizontal lines indicate medians; whiskers denote range; and outliers are shown as individual points. To note, the expert-made performance checklist showed strong internal consistency ($\alpha = 0.83$) and good item discrimination ($D = 0.60-1.00$), indicating reliable assessment of procedural performance across varying levels of task difficulty.

A secondary outcome, confidence ratings (3-point scale, Figure 3), differed significantly by group for observer raters ($p = 0.019$), with lower scores in the AI group, while SP ratings showed no significant differences ($p = 0.243$); observer ratings strongly correlated with performance ($r = 0.84$, $p < 0.001$). Of note, the observer ratings differed from the ratings of SPs (Figure 3).

How to cite this article:

Palaganas J C, Morton A, Adem F, et al. (June 02, 2026) Assessment of Student Learning Outcomes Using AI and Simulation in Choking Management Skills: A Pilot Randomized Controlled Trial. *Cureus J Comput Sci* 3 : es44389-026-00094-y. DOI <https://doi.org/10.7759/s44389-026-00094-y>

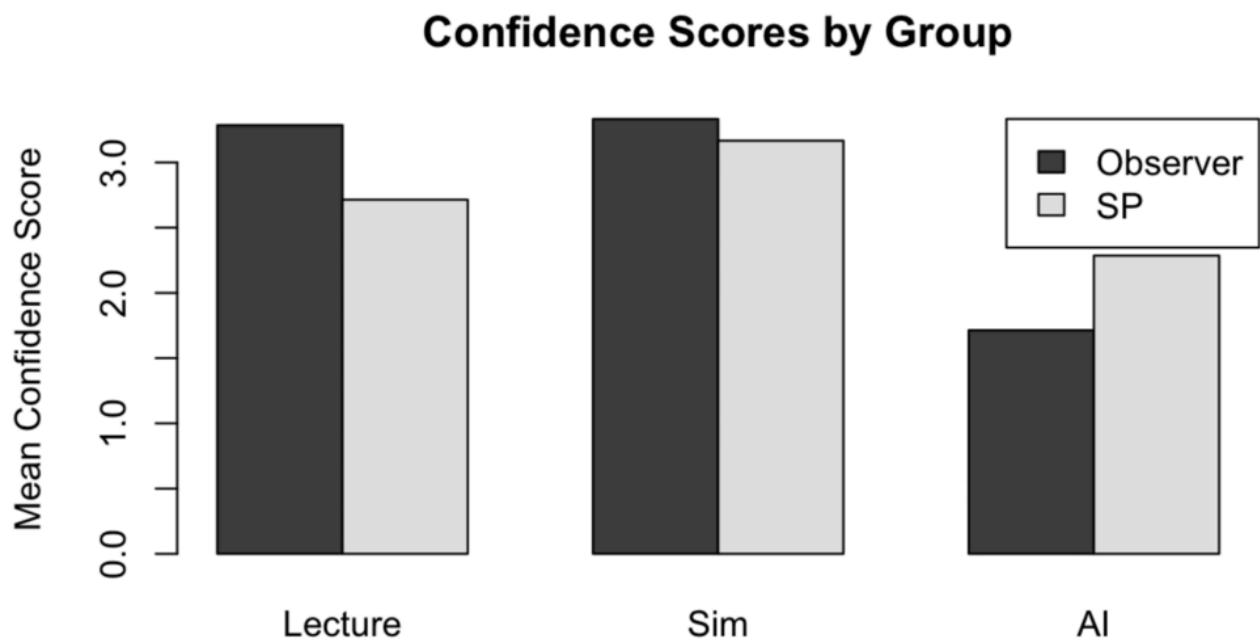


FIGURE 3: Confidence Score Comparison
Grouped barplot depicting mean observer and standardized patient (SP) confidence scores by instructional group (Lecture, Simulation, and AI-assisted). Dark gray bars represent observer confidence scores; light gray bars represent SP confidence scores. Each instructional group displays side-by-side comparisons of mean confidence ratings between rater types. Error bars not shown.

Given the small individual group sizes ($n = 6-7$), we used change score analysis to maximize statistical power for detecting differences in knowledge retention patterns between teaching methods [23,24]. Overall, students showed minimal knowledge decline from post-intervention to follow-up ($M = -0.18$ points, 95% CI $[-1.41, 1.06]$, $p = 0.770$). There was a marginally significant difference between groups in retention patterns ($F = 3.06$, $p = 0.073$), suggesting that teaching method may differentially affect knowledge retention over time; given the underpowered sample, we treat this pattern as hypothesis-generating rather than confirmatory.

Regression Analysis

Simulation yielded statistically higher knowledge scores than Lecture at both post-test and follow-up, while AI was associated with lower performance, and male students consistently scored lower than female students in knowledge and critical thinking; no effects were observed for age, race, or grade (Table 2). We note that, given the small sample ($N = 20$), the multivariable regression models include more predictors than the data can support with confidence; we report these associations as preliminary and hypothesis-generating, not as confirmatory effects, and we do not draw causal inferences from them.

Predictor	Total Knowledge (Post)	Knowledge (Follow-up)	Performance (Action)	Critical Thinking (Post)	Critical Thinking (Follow-up)
Lecture(ref)	-	-	-	-	-
Sim	3.74 (2.06) *	4.58 (1.85) **	-0.21 (1.32)	0.30 (0.48)	0.17 (0.59)
AI	-3.73 (2.08)	0.06 (1.87)	-3.64 (1.33) **	-0.65 (0.48)	-0.26 (0.59)
Covariates					
Age	1.78 (3.87)	2.16 (3.48)	-0.82 (2.48)	0.15 (0.90)	0.16 (1.10)
Male (vs. Female)	-4.16 (1.86) **	-3.23 (1.67) *	0.66 (1.19)	-0.72 (0.43)	-1.44 (0.53) **
Non-binary	12.10 (8.42)	12.99 (7.56)	-2.53 (5.41)	3.23 (1.95)	0.24 (2.40)
Hispanic	-0.92 (2.58)	-0.75 (2.32)	0.49 (1.66)	0.78 (0.60)	-1.09 (0.73)
Multi-race	0.18 (2.37)	-0.75 (2.13)	1.69 (1.52)	-0.38 (0.55)	-0.82 (0.68)
White	2.47 (1.64)	0.36 (1.48)	0.61 (1.06)	0.40 (0.38)	-0.16 (0.47)
Grade	-4.35 (3.83)	-4.80 (3.44)	0.89 (2.46)	-1.13 (0.89)	-0.80 (1.09)
Constant	48.39 (25.73) *	46.22 (23.11) *	10.16 (16.52)	12.09 (5.95) *	9.98 (7.33)
Model Statistics					
R ²	0.744	0.648	0.655	0.774	0.575
Adj. R ²	0.514	0.332	0.345	0.570	0.192
F(9,10)	3.24 **	2.05	2.11	3.80 **	1.50

TABLE 2: Regression Analysis Outcomes

Standard errors in parentheses. * $p < .10$; ** $p < .05$; *** $p < .01$. Assumptions were assessed via Shapiro–Wilk and Levene’s tests. ANOVA was used for variables meeting parametric criteria (e.g., Total Knowledge, Grade); Kruskal–Wallis was used for others. Regression analyses, controlled for demographics, explained moderate to high variance ($R^2 = 0.58–0.77$).

AI Engagement Patterns

Analysis of the AI-assisted group’s ChatGPT transcripts revealed substantial variability in how learners enacted the intervention (consistent with the variability noted across all participants who had access to AI during the assessment). All AI-group students submitted complete chat threads. Several patterns emerged. First, prompt strategy varied widely: most students copied test or outline items verbatim into the chat without reformulation, while a smaller subset progressively refined prompts, asked follow-up clarifying questions, or pushed back on AI responses. Second, students

How to cite this article:

Palaganas J C, Morton A, Adem F, et al. (June 02, 2026) Assessment of Student Learning Outcomes Using AI and Simulation in Choking Management Skills: A Pilot Randomized Controlled Trial. *Cureus J Comput Sci* 3 : es44389-026-00094-y. DOI <https://doi.org/10.7759/s44389-026-00094-y>

differed in how they incorporated AI output: some synthesized AI content with their own knowledge, others rephrased AI outputs in their own words, and a few used minimal or no AI content in their final answers. Third, students differed in fact-checking behavior: some fact-checked every AI response, some only when the response felt uncertain, and others accepted initial outputs without further scrutiny. Fourth, a subset of students explicitly requested academic references from the AI; this was associated with longer, more elaborated chats. Fifth, several students reported feelings of guilt or performance bias around AI use (also captured in the focus group, below), which appeared to shape how much they engaged with the tool during the assessment. We pull these patterns into the Results so that readers can see how the AI “intervention” was actually enacted by participants; we discuss their implications in the Discussion.

Focus Group

We found two thematic categories: 1) Helpful/Not Helpful and 2) Confounding Psychological Factors. Overall, the group found AI to be helpful, some students stating, “It definitely helped with the essay questions” and “It definitely helped me with one multiple choice question.” While few students felt that AI was not helpful, some stating, “I already knew the answers” or “It kept generating the same thing and wasn’t that helpful.”

The conversation revealed confounding psychological factors including guilt and performance bias:

Guilt: “I felt like I was cheating so I really didn’t want to use it, but I did just because I knew it was part of the research.”

Performance Bias: “I knew you were going to be looking at the chat threads, so I just copied and pasted a few questions even though I didn’t really need to use it.”

The group also discussed the need to better learn prompt design. All students agreed that they “simply copied and pasted” the test questions into ChatGPT. The group discussed that if they learned prompt design before using AI to support test-taking, it might have improved their use.

Discussion

The findings of this pilot study offer preliminary insights into how AI-assisted self-directed learning may compare with traditional lecture-based instruction and simulation-based education within pre-professional training. While AI tools provided students with flexible, self-directed opportunities to explore course content, the data are suggestive of distinctions in performance, confidence, and critical thinking across the three arms. We emphasize, however, that the design as implemented contrasted faculty-supported teaching (lecture and simulation) with largely unguided self-directed study using AI, rather than AI-assisted teaching. Interpreting these outcomes in context appears to highlight both the continued strengths of faculty-guided learning environments and the emerging opportunities-and current limitations-of integrating AI into early clinical education.

AI-Assisted Self-Learning vs Lecture vs Simulation-Based Education

Evaluating the impact of AI-assisted learning compared to traditional lecture and simulation methods within pre-professional student education reveals distinct patterns in clinical performance, competency, critical thinking, and knowledge retention.

Clinical Performance and Competency: The data appear to suggest that lecture- and simulation-based education (SBE) groups performed similarly and better than AI-assisted self-learning (the “AI group”). This may indicate that faculty-driven sessions provide an advantage for developing hands-on skills and meeting competency benchmarks, although in this study, the AI arm also received less instructional support and no instruction on best use of AI, which complicates direct attribution to the modality itself. It is noteworthy that the SBE group performed slightly better than the lecture-based method, but this difference was not statistically significant, which is consistent with both methods being potentially useful when paired with faculty engagement.

Confidence: Confidence levels mirrored the performance pattern, with participants in lecture-based and SBE groups demonstrating more confident movement and communication when compared with the AI group. This increased confidence may stem from the personalized and immediate feedback that faculty provided in those arms. It may also

How to cite this article:

Palaganas J C, Morton A, Adem F, et al. (June 02, 2026) Assessment of Student Learning Outcomes Using AI and Simulation in Choking Management Skills: A Pilot Randomized Controlled Trial. *Cureus J Comput Sci* 3 : es44389-026-00094-y. DOI <https://doi.org/10.7759/s44389-026-00094-y>

reflect that the AI group was not provided any faculty-chosen learning material (e.g., reading or video) beyond the course outline prior to AI use, leaving them less anchored in a curated knowledge base before the assessment.

Confidence serves as a key factor in strengthening the link between medical training and clinical performance, with higher confidence driving more favorable outcomes [25]. In this study, confidence is a crucial metric when evaluating if AI, lecture-based, or simulation models translate into real-world preparedness and clinical competency. Evaluating which educational method produces confident learners ensures the study captures not only knowledge acquisition but also the development of practical clinical effectiveness and patient care.

Knowledge Test Results: The lecture-based group performed strongly on multiple-choice assessments, which is consistent with a good grasp of factual and procedural knowledge bolstered by AI. The SBE group, however, performed strongly on essay responses, which may indicate facilitation of deeper conceptual thinking and synthesis skills with AI support. Both lecture-based and SBE groups outperformed the AI group overall, again suggestive of an advantage for faculty support in enhancing knowledge acquisition, critical thinking, confidence, and performance, regardless of assessment format, although we underscore that this pattern is consistent with the asymmetry of instructional support.

Critical Thinking Development: Analysis of AI chat threads suggested that the lecture-based group demonstrated the highest level of critical thinking when interacting with AI, with the SBE group slightly behind. Both the lecture-based and the SBE groups appeared ahead of the AI group in critical thinking when scored with the modified HCTSR. This pattern is consistent with a reading in which faculty-led learning helps prime the cognitive processes useful for clinical reasoning and judgment.

Feasibility of Assessing AI-Generated Chat Transcripts for Knowledge and Critical Thinking

The process of evaluating AI-generated chat transcripts for knowledge and critical thinking proved to be surprisingly feasible, especially when applying a modified version of the HCTSR. While the rubric's inherent subjectivity and the complex constructs of critical thinking necessitated a consensus approach among raters, it was still possible to discern meaningful patterns and assess student performance effectively. Some students approached the transcript questions in order, demonstrating a linear progression in thought. Others skipped and returned to difficult questions, reflecting metacognitive flexibility. Copying questions word-for-word was a frequent strategy, showing direct engagement. Certain students fact-checked every question using AI, seeking comprehensive accuracy. Others used AI only when unclear, suggesting varied levels of confidence and skill with the technology. Some students refined prompts, asked for clarifications, and challenged AI responses, illustrating critical engagement. Others accepted initial AI outputs without further inquiry, displaying a more passive approach. A portion of students requested academic references, emphasizing the importance of credible sourcing. And, finally, when constructing final answers, some students synthesized AI-generated information with their own knowledge, creating integrated responses. Others rephrased AI outputs in their own words, demonstrating transformative learning. A few made only minimal changes or composed entirely original answers, while some chose to leave out AI content altogether. Rewording or breaking questions into smaller components indicated deeper processing and active understanding.

Feelings of guilt and performance bias influenced how students engaged with AI, shaping both learning behavior and assessment authenticity. Guilt reflected moral tension in AI use, while performance bias revealed self-presentation concerns. These findings imply that AI-integrated assessments must address learner psychology in its use.

Advantages and Challenges of Using Simulation Performance and Knowledge Test and AI Chat Threads in Assessing Student Learning Outcomes

Integrating simulation performance, knowledge tests, and AI chat threads offers a multidimensional view of learning outcomes by linking observable behaviors, factual knowledge, and reasoning processes. This triangulation enhances findings, revealing both what students know, how they think, and how they apply knowledge. However, challenges include ensuring interrater reliability, managing subjective interpretation of AI interactions, and balancing technology's benefits with ethical and logistical concerns surrounding data privacy, scoring consistency, and equitable access to AI tools.

How to cite this article:

Palaganas J C, Morton A, Adem F, et al. (June 02, 2026) Assessment of Student Learning Outcomes Using AI and Simulation in Choking Management Skills: A Pilot Randomized Controlled Trial. *Cureus J Comput Sci* 3 : es44389-026-00094-y. DOI <https://doi.org/10.7759/s44389-026-00094-y>

Strengths and Limitations

This study is among the first to integrate generative AI-based self-learning with simulation and lecture methods to assess pre-professional students' clinical and critical thinking skills. Combining performance assessments, knowledge tests, and AI chat analyses provided a comprehensive view of learning outcomes. Reliable tools, authentic simulations, and consensus-based scoring enhanced validity, while studying technologically proficient students offered insights to guide future AI-integrated health professions education.

This was a single-site pilot study, and therefore, the findings should be interpreted as preliminary and hypothesis-generating. Several limitations are important for interpretation. First, by design, the AI-assisted group received considerably less instructional support than the lecture and simulation groups: there was no faculty instruction, AI literacy scaffolding, or curated reading or video materials beyond the standardized outline. In effect, the comparison may reflect unguided self-study with AI versus faculty-supported teaching as much as it reflects AI as a modality. Second, the AI group differed demographically from the other arms by chance of randomization (i.e., slightly older, entirely female, more racially diverse, and with a higher proportion of EAL students), which may have contributed to the lower performance and confidence scores independent of instructional modality. Third, no baseline knowledge or critical-thinking assessment was administered, which limits inferences about change attributable to instruction. Fourth, the multivariable regression models include more predictors than the small sample can robustly support; we therefore report those associations as suggestive rather than confirmatory. Fifth, some ChatGPT interactions were lost when prompts were replaced with "<content removed>", limiting review of design. Sixth, camera angles and occasional facilitator prompting affected scoring, and variations in communication style influenced scoring fidelity. Finally, several limitations are specific to the LLM intervention itself: ChatGPT-4o was used as an off-the-shelf general-purpose conversational model rather than a model trained or fine-tuned on emergency-care or pediatric resuscitation curricula; its outputs are sensitive to prompt phrasing, and learners in this study were not coached in prompt design before use; the chat interface itself favors text-based engagement and does not natively include integrated demonstrations, structured feedback loops, or spaced retrieval practice. Future implementations might pair general-purpose LLMs with curriculum-specific scaffolds (e.g., prompt libraries, retrieval-augmented conditioning of LLMs on current AHA/PALS guidelines, or instructor-mediated chat reviews) to better isolate the contribution of AI as a modality. We note that small sample size is, of course, expected for a pilot study; we list it here because the underpowered sample affects the precision of effect estimates rather than because it is a defect of the design. Future studies should include baseline assessment, larger and more diverse cohorts, comparable instructional support across arms, explicit AI-literacy and prompt-design scaffolding, and prospective registration.

Future Directions

This pilot opens several specific directions for future work. Future studies should equalize instructional support across arms, either by adding faculty-mediated AI use (e.g., a brief AI-literacy and prompt-design primer, or facilitator-reviewed chats) or by comparing different forms of AI scaffolding head-to-head; this would allow more direct attribution of differences to the modality rather than to the level of support. Larger, multi-site samples with baseline assessment of knowledge and critical thinking, stratified or blocked randomization on demographic and language variables (including English as an additional language status), and prospective registration would improve internal validity and generalizability. AI literacy itself deserves study as an outcome and as a mediator: how does explicit instruction in prompt engineering, source-checking, and critical evaluation of AI output shape learning gains in clinical content domains? The psychological dimensions surfaced in our focus group (i.e., guilt, performance bias, and uncertainty about appropriate use) warrant a dedicated study; learners need clear norms for AI use in clinical training, and educators need assessment formats that account for those norms. Blended designs that integrate AI with simulation and lecture (rather than treating them as competitors) are likely the practically important next step; future work should examine, for example, AI-mediated pre-briefing, AI-supported debriefing aids, and AI-assisted reflection after simulation. Technical evolution of the LLM itself matters: studies should examine whether retrieval-augmented or curriculum-tuned chat agents (e.g., grounded in current AHA/PALS guidelines and HSSOBPTM) outperform general-purpose models in supporting clinical learning, and

How to cite this article:

Palaganas J C, Morton A, Adem F, et al. (June 02, 2026) Assessment of Student Learning Outcomes Using AI and Simulation in Choking Management Skills: A Pilot Randomized Controlled Trial. *Cureus J Comput Sci* 3 : es44389-026-00094-y. DOI <https://doi.org/10.7759/s44389-026-00094-y>

whether multimodal interfaces (voice, video, integrated demonstration) change the engagement patterns we observed. Together, these directions suggest a research agenda that treats AI not as a drop-in substitute for faculty teaching but as a complementary, increasingly fine-tunable component of health professions education.

Conclusions

In this small pilot, lecture- and simulation-based instruction appeared to produce higher clinical performance, confidence, and critical thinking than AI-assisted self-learning among pre-professional students. The pattern is consistent with prior literature suggesting that guided and feedback-rich learning environments can support competency and reasoning more effectively than unstructured self-directed digital learning; we caution, however, that in this study the AI group also received less instructional support than the comparator arms, so these findings should be interpreted as preliminary and hypothesis-generating rather than as a definitive comparison of AI as a modality. Even with that caveat, the data are suggestive that early development of AI literacy, prompting competency, and psychological safety may matter for successful integration of AI into health professions education. Future research should expand sample diversity, pair AI with comparable faculty-supported scaffolding, standardize procedures, and further explore how AI transcript analysis can enhance the assessment of reasoning and metacognition. Overall, this study demonstrates the feasibility and the limits of integrating AI and simulation in performance-based assessment in health professions education and points to specific, testable directions for future research.

Appendices

Appendix I. Simulation Case on Choking

Infant Choking Simulation Case

EQUIPMENT AND SETUP

- Infant manikin capable of stridor and crying sounds
- Bags of grapes or prop food
- Chair
- Stethoscope
- Bedside table
- History form and pen

Case Stem

We are in a primary care clinic. You are Medical or Nursing Students observing for the day. Your clinical advisor or preceptor has asked you to start your observations by going in first and getting a history on all of the patients. Here is a history sheet that your care team will need to fill out to give to your clinical preceptor.

Roles

- Learners: Medical/ Nursing Students
- Standardized Patient: Patient Mom
- Embedded Simulation Person (ESP):Nurse

History Form Template Provided to Learners

- Date of birth:
- What brings you in today?

How to cite this article:

Palaganas J C, Morton A, Adem F, et al. (June 02, 2026) Assessment of Student Learning Outcomes Using AI and Simulation in Choking Management Skills: A Pilot Randomized Controlled Trial. *Cureus J Comput Sci* 3 : es44389-026-00094-y. DOI <https://doi.org/10.7759/s44389-026-00094-y>

- OLDCART Assessment
 - Onset:
 - Location:
 - Duration:
 - Characteristics (Can you describe what it feels like?):
 - Aggravating Factors (What makes it feel worse?):
 - Relieving Factors (What makes it feel better?):
 - Treatment (Are there any medicines that you're taking that helps?):

SCENARIO PROGRESSION

Initial Stage

Outside of simulation room, learners are told:

"Mrs. Brennan is here since she's been having allergies. She's in room 4 waiting for you."

Upon entry, the patient (Amy Brennan) is sitting in a chair holding her 8-month-old infant, Annie, and a small container of grapes.

Patient Information for Standardized Patient

- Name: Amy Brennan
- DOB: 08/01/1993 (32 years old)
- Infant: Annie, 8 months old
- Chief Complaint: "I've been having allergies every night and every morning."

Sample History Responses

- Onset: "I think it started after I had Annie here (8 month-old)."
- Location: "My nose gets runny, my throat gets itchy, sometimes my lips turn red and my skin gets little bumps and is itchy."
- Duration: "It lasts for about 2 hours in the morning and an hour before bed."
- Characteristics: "Like a runny nose and I feel like I can't breathe..."
 - (At this point, the parent notices the infant is not breathing.)

Second Stage (Trigger Event)

- Cue from Manikin: [cough followed by stridor sound]
- Cue from Standardized Patient: "ANNIE, OMG, ANNIE?! ANNIE!" SHE'S BLUE. SHE'S NOT BREATHING! WAIT I THINK SHE'S CHOKING!!! OMG OMG OMG!! HELP ME!! OMG.

Critical Actions Checklist: Expected Actions from Learners

- Supports infant's head and neck
- Delivers 5 back blows between shoulder blades
- Turns infant over and gives 5 chest thrusts
- Visually Checks mouth for obstruction between attempts

How to cite this article:

Palaganas J C, Morton A, Adem F, et al. (June 02, 2026) Assessment of Student Learning Outcomes Using AI and Simulation in Choking Management Skills: A Pilot Randomized Controlled Trial. Cureus J Comput Sci 3 : es44389-026-00094-y. DOI <https://doi.org/10.7759/s44389-026-00094-y>

Final Stage

- Cue from Manikin: [crying]
- Cue from Standardized Patient: OMG THANK YOU THANK YOU!!!
- Cue from ESP Nurse
 - Dialogue: "What happened in here?! I'm going to take over from here. Thank you, you can go back to your room."

Appendix II. Choking Performance Measure Tool

How to cite this article:

Palaganas J C, Morton A, Adem F, et al. (June 02, 2026) Assessment of Student Learning Outcomes Using AI and Simulation in Choking Management Skills: A Pilot Randomized Controlled Trial. Cureus J Comput Sci 3 : es44389-026-00094-y. DOI <https://doi.org/10.7759/s44389-026-00094-y>

	Description of the Action	Assessment Score From 0-1 (Did/Did Not Do)
1	Provider reached for the baby	
2	Response time (time from acquiring the baby to back blows in seconds)	
3	Support of the infant's head and neck	
4	Placement of hands for back blows	
5	Delivered 5 back blows	
6	Placement of hands for chest thrusts	
7	Delivered 5 chest thrusts	
8	Checked airway/Looked inside mouth	
9	Repeated cycle	
	Total (do not include (#2) response time)	
10	Overall likert score on performance (1=not good, 2=good, 3=excellent)	
11	Explain your overall rating of performance:	
12	Overall likert score on confidence (1=not good e.g. microexpressions seemed lost, 2=okay, e.g. moving confidently but not speaking confidently or vice versa, 3=good, e.g., mostly moving confidently but not speaking confidently or vice versa, 4=excellent, e.g. communicated confidently and more than once)	
13	Explain your overall rating on confidence:	

TABLE 3: Choking Performance Measure Tool

Appendix III. Knowledge and Essay Test

UNIQUE ID:

How to cite this article:

Palaganas J C, Morton A, Adem F, et al. (June 02, 2026) Assessment of Student Learning Outcomes Using AI and Simulation in Choking Management Skills: A Pilot Randomized Controlled Trial. Cureus J Comput Sci 3 : es44389-026-00094-y. DOI <https://doi.org/10.7759/s44389-026-00094-y>

(Father's initials)-(favorite color)-(house number on street address)

DATE:

Choking Response Test

Instructions: Answer all multiple-choice and essay questions.

Multiple Choice

Which sign indicates a complete airway obstruction in a conscious adult? Answer:

- a) Forceful coughing
- b) Clutching the throat (universal choking sign)
- c) Wheezing during inhalation
- d) Complaining of chest pain

The correct hand placement for abdominal thrusts in a non-pregnant adult is: Answer:

- a) Just below the ribcage
- b) Over the sternum
- c) Slightly above the navel
- d) Midway between the hips and ribs

For a choking infant (<1 year), which sequence is correct? Answer:

- a) 5 chest thrusts → 5 back blows → repeat
- b) 5 back blows → 5 chest thrusts → repeat
- c) 5 abdominal thrusts → 5 back blows → repeat
- d) 5 back blows → 5 abdominal thrusts → repeat

A key prevention strategy for choking in toddlers is: Answer:

- a) Serving whole grapes and nuts
- b) Cutting food into ½-inch pieces
- c) Allowing unsupervised play with small toys
- d) Encouraging eating while running

If a choking person becomes unconscious, the first step is: Answer:

- a) Perform abdominal thrusts
- b) Begin CPR with chest compressions
- c) Sweep the mouth for obstructions
- d) Call 911

Which is an example of a mild foreign-body airway obstruction? Answer:

- a) cyanosis (blue lips or skin)
- b) high pitched noise while inhaling
- c) inability to speak or cry
- d) wheezing between cough

Which victim of a severe airway obstruction should receive abdominal thrusts? Answer:

- a) an average-size 27-year-old man
- b) a woman who is obviously pregnant
- c) an obese 50-year-old man
- d) an average-size 9-month-old infant

You are performing abdominal thrusts on a 9-year-old child when she suddenly becomes unresponsive. After you seek nearby help, what is the most appropriate action to take next? Answer:

How to cite this article:

Palaganas J C, Morton A, Adem F, et al. (June 02, 2026) Assessment of Student Learning Outcomes Using AI and Simulation in Choking Management Skills: A Pilot Randomized Controlled Trial. Cureus J Comput Sci 3 : es44389-026-00094-y. DOI <https://doi.org/10.7759/s44389-026-00094-y>

- a) Begin high-quality CPR, starting with chest compressions
- b) Check for a pulse
- c) Continue performing abdominal thrusts
- d) Provide 5 back slaps followed by 5 chest thrusts

I should perform a blind finger sweep if I suspect that there is a foreign body in the airway: True/False

Short Answer

- List three foods that pose high choking risks for children under 4 and explain how to prepare them safely.
- Describe the difference in technique when assisting a choking pregnant individual versus a non-pregnant adult.
- Why should abdominal thrusts never be used on infants? What alternative maneuvers are recommended?

Short Answer Essay Questions

1. A 7-year-old child is choking on a marble at a park. They are conscious but unable to cough or speak.

Outline the steps you would take to assist them.

Explain how you would modify your approach if the child became unconscious.

2. Design a community education initiative to reduce choking incidents in daycare centers.

Identify three target behaviors for staff/parents.

Propose two hands-on training activities.

Discuss how to address cultural dietary practices that may increase risk.

3. An elderly patient with Parkinson's disease frequently chokes on meals.

Analyze three risk factors specific to this patient.

Recommend evidence-based strategies to improve safety during mealtimes.

ESSAY QUESTIONS

4. How would your response differ between a choking adult, child, and infant, and why is it important to tailor your actions to the person's age and size?

5. If the person becomes unresponsive while you're providing care, what is your next step, and what critical cues help you decide when to switch from choking intervention to CPR?

6. What are some signs that a person is experiencing a mild airway obstruction versus a severe one, and how might acting too soon or too late in either case affect the outcome?

Appendix IV. Modified for AI Holistic Critical Thinking Scoring Rubric

How to cite this article:

Palaganas J C, Morton A, Adem F, et al. (June 02, 2026) Assessment of Student Learning Outcomes Using AI and Simulation in Choking Management Skills: A Pilot Randomized Controlled Trial. *Cureus J Comput Sci* 3 : es44389-026-00094-y. DOI <https://doi.org/10.7759/s44389-026-00094-y>

Level	Holistic Description	How It Presents in AI Essays/Chat Threads
Level 4: Advancing	Consistently does all or almost all of the following: Accurately interprets evidence, statements, graphics, questions, and so on. Identifies the salient arguments (reasons and claims), pros, and cons. Thoughtfully analyzes and evaluates major alternative points of view. Draws warranted, judicious, and non-fallacious conclusions. Justifies key results and procedures, and explains assumptions and reasons. Fair-mindedly follows where evidence and reasons lead	Consistently (at least 90% of their thread): Interprets AI responses correctly, including ignoring or questioning inaccurate data. Provides a detailed accurate synthesis of what they discovered from AI, acknowledges pros and cons if applicable, and used previous AI responses in other questions. Asks additional prompts, including clarifying prompts #1 and #2 accurately. Consistently provides accurate and detailed reasoning. Organized evidence-based arguments. Deep thought (very much student voice)
Level 3: Achieving	Does most or many of the following: Accurately interprets evidence, statements, graphics, questions, and so on. Identifies relevant arguments (reasons and claims), pros, and cons. Offers analyses and evaluations of obvious alternative points of view. Draws warranted, nonfallacious conclusions. Justifies some results or procedures, and explains reasons. Fair-mindedly follows where evidence and reasons lead	Mostly (less than 90% of their thread): Interprets AI responses correctly. Provides a clear synthesis of what they discovered from AI. Acknowledges pros and cons if applicable. Asks additional prompts #1 and #2 accurately. Consistently provides reasoning. Organized evidence-based arguments. Hear student voice, but not entering deep thought, and not asking clarification questions of AI
Level 2: Developing	Does most or many of the following: Misinterprets evidence, statements, graphics, questions, and so on Fails to identify strong, relevant counterarguments Ignores or superficially evaluates obvious alternative points of view Draws unwarranted or fallacious conclusions Justifies few results or procedures, and seldom explains reasons Regardless of the evidence or reasons, maintains or defends views based on self-interest or preconceptions	Mostly (less than 90% of their thread): Misinterprets question or correct AI responses Superficial synthesis, accurate answers with no support Only copies and pastes test question Enters inaccurate answers Provides little to no reasons #2. Partially incorrectly or AI is doing most of thinking for them (do not hear their voice)
Level 1: Initiating	Consistently does all or almost all of the following: Offers biased interpretations of evidence, statements, graphics, questions, information, or the points of view of others. Fails to identify or hastily dismisses strong, relevant counterarguments. Ignores or superficially evaluates obvious	Consistently (at least 90% of their thread): In test essay, there is inaccurate data, ignores accurate AI info, misinterprets AI answer to fit their own bias, and does not ask AI or only copies and pastes test question. Uses prompts that argue with AI when AI answer is accurate and enters

How to cite this article:

Palaganas J C, Morton A, Adem F, et al. (June 02, 2026) Assessment of Student Learning Outcomes Using AI and Simulation in Choking Management Skills: A Pilot Randomized Controlled Trial. Cureus J Comput Sci 3 : es44389-026-00094-y. DOI <https://doi.org/10.7759/s44389-026-00094-y>

	alternative points of view. Argues using fallacious or irrelevant reasons and unwarranted claims. Does not justify results or procedures, and fails to explain reasons. Regardless of the evidence or reasons, maintains or defends views based on self-interest or preconceptions. Exhibits closed-mindedness or hostility to reason	inaccurate answers. Fails to provide reason #2 All of the above. AI is doing all of the thinking or getting wrong answer
--	---	--

TABLE 4: Modified for AI Holistic Critical Thinking Scoring Rubric

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Janice C. Palaganas, Alex Morton, Khaulah Jawed

Acquisition, analysis, or interpretation of data: Janice C. Palaganas, Alex Morton, Fatima Adem, Jianna Ramos, Arjun Kumar, Maria Bajwa

Drafting of the manuscript: Janice C. Palaganas, Alex Morton, Fatima Adem, Jianna Ramos, Arjun Kumar, Khaulah Jawed, Maria Bajwa

Critical review of the manuscript for important intellectual content: Janice C. Palaganas, Alex Morton, Jianna Ramos, Arjun Kumar

Supervision: Janice C. Palaganas

Disclosures

Human subjects: Consent was obtained or waived by all participants in this study. MGB Institutional Review Board issued approval 2025P001575. This study was deemed exempt by the MGB Institutional Review Board. **Animal subjects:** All authors have confirmed that this study did not involve animal subjects or tissue. **Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

Data Availability Statements

The datasets (and/or code) supporting this study are available from the corresponding author upon reasonable request.

Acknowledgements

The authors would like to thank Dr. Anshul Kumar for his review and statistical guidance and Dr. Suzie Kardong-Edgren for her thoughts and feedback on our work.

How to cite this article:

Palaganas J C, Morton A, Adem F, et al. (June 02, 2026) Assessment of Student Learning Outcomes Using AI and Simulation in Choking Management Skills: A Pilot Randomized Controlled Trial. Cureus J Comput Sci 3 : es44389-026-00094-y. DOI <https://doi.org/10.7759/s44389-026-00094-y>

References

1. Clusmann J, Kolbinger FR, Muti HS, et al.: [The future landscape of large language models in medicine](#). Communications Medicine. 2023, 3:141. [10.1038/s43856-023-00370-1](#)
2. Maaz S, Palaganas JC, Palaganas G, Bajwa M: [A guide to prompt design: foundations and applications for healthcare simulationists](#). Frontiers in Medicine. 2025, 11:1504532. [10.3389/fmed.2024.1504532](#)
3. Kouka A, Giannelou E, Konstantinidis K, Apostolakis I: [Implementation of artificial intelligence in nursing education: A narrative review](#). Health & Research Journal. 2025, 11:74-82. [10.12681/healthresj.36795](#)
4. Sajja R, Sermet Y, Cikmaz M, Cwiertny D, Demir I: [Artificial intelligence-enabled intelligent assistant for personalized and adaptive learning in higher education](#). Information. 2024, 15:596. [10.3390/info15100596](#)
5. Ramírez-Baraldes E, García-Gutiérrez D, García-Salido C: [Artificial intelligence in nursing: new opportunities and challenges](#). European Journal of Education. 2025, 60:e70033. [10.1111/ejed.70033](#)
6. Topaz M, Peltonen L-M, Michalowski M, et al.: [The ChatGPT effect: nursing education and generative artificial intelligence](#). Journal of Nursing Education. 2025, 64:e40-e43. [10.3928/01484834-20240126-01](#)
7. Vieriu AM, Petrea G: [The impact of artificial intelligence \(AI\) on students' academic development](#). Education Sciences. 2025, 15:343. [10.3390/educsci15030343](#)
8. Shibani A, Knight S, Kitto K, Karunanayake A, Buckingham Shum S: [Untangling critical interaction with AI in students' written assessment](#). CHI EA '24: Extended Abstracts of the CHI Conference on Human Factors in Computing Systems. 2024, 1-6. [10.1145/3613905.3651083](#)
9. Gilbert M, Watts P, Brown KM: [Competency-based education: Simulation educators, don't be Afraid!](#). Clinical Simulation In Nursing. 2024, 97:101668.
10. Ong H, Ong J, Chen D, Li S, Ong D: [A novel artificial intelligence leadership curriculum for pre-college students interested in medicine and engineering: program development and global competition outcomes](#). Journal of Medical Artificial Intelligence. 2025, 8:1-13.
11. Lawasi MC, Rohman VA, Shoreamanis M: [The use of AI in improving student's critical thinking skills](#). Proceedings Series on Social Sciences & Humanities. 2024, 18:366-370. [10.30595/pssh.v18i.1279](#)
12. Facione PA, Facione NC: [The Holistic Critical Thinking Scoring Rubric \[Internet\]](#). <https://insightassessment.com/iaresource/the-holistic-critical-thinking-scoring-rubric/>.
13. Facione PA, Facione NC: [Critical Thinking and Clinical Reasoning in the Health Sciences: An International Multidisciplinary Teaching Anthology](#). Insight Assessment/The California Academic Press, Millbrae, CA; 2008.
14. Kendig CE: [What is proof of concept research and how does it generate epistemic and ethical categories for future scientific practice?](#). Science and Engineering Ethics. 2016, 22:735-753. [10.1007/s11948-015-9654-0](#)
15. Watts PI, Rossler K, Bowler F, et al.: [Onward and upward: introducing the healthcare simulation standards of best practiceTM](#). Clinical Simulation in Nursing. 2021, 58:1-4. [10.1016/j.ecns.2021.08.006](#)
16. ChatGPT [Internet]. <https://chatgpt.com/>.
17. Lasa JJ, Dhillion GS, Duff JP, et al.: [Part 8: Pediatric advanced life support: 2025 American Heart Association and American Academy of Pediatrics guidelines for cardiopulmonary resuscitation and emergency cardiovascular care](#). Circulation. 2025, 152:S479-S573. [10.1161/cir.0000000000001368](#)
18. Wigginton JG, Agarwal S, Bartos JA, et al.: [Part 9: adult advanced life support: 2025 American Heart Association guidelines for cardiopulmonary resuscitation and emergency cardiovascular care](#). Circulation. 2025, 152:S538-S577. [10.1161/cir.0000000000001376](#)
19. Zlatkin-Troitschanskaia O, Shavelson RJ, Schmidt S, Beck K: [On the complementarity of holistic and analytic approaches to performance assessment scoring](#). British Journal of Educational Psychology. 2019, 89:468-484. [10.1111/bjep.12286](#)
20. Layman HA: [A Robust and Generalisable Rubric Design Framework for Critical Thinking Assessment](#). University of Leicester. Thesis. (2021). <https://doi.org/10.25392/leicester.data.14745837.v1>.
21. Dedoose. (2025). <https://www.dedoose.com>.

How to cite this article:

Palaganas J C, Morton A, Adem F, et al. (June 02, 2026) Assessment of Student Learning Outcomes Using AI and Simulation in Choking Management Skills: A Pilot Randomized Controlled Trial. Cureus J Comput Sci 3 : es44389-026-00094-y. DOI <https://doi.org/10.7759/s44389-026-00094-y>

22. Schreier M: [Qualitative Content Analysis in Practice](#). Sage, Cambridge; 2012.
23. Hedges LV, Rhoads C: [Statistical Power Analysis in Education Research](#). National Center for Special Education Research, Washington, DC; 2010.
24. Estrada E, Ferrer E, Pardo A: [Statistics for evaluating pre-post change: relation between change in the distribution center and change in the individual scores](#). Frontiers in Psychology. 2019, 9:2696. [10.3389/fpsyg.2018.02696](#)
25. Yazbeck Karam VG, Bou Malhab S, Aoun Bahous S, Salameh P, El Khoury-Malhame M, Asmar N: [The mediating role of self-confidence in medical students' attitudes and clinical performance toward people with disability](#). Advances in Medical Education and Practice. 2025, 16:1169-1179. [10.2147/AMEP.S522243](#)

How to cite this article:

Palaganas J C, Morton A, Adem F, et al. (June 02, 2026) Assessment of Student Learning Outcomes Using AI and Simulation in Choking Management Skills: A Pilot Randomized Controlled Trial. Cureus J Comput Sci 3 : es44389-026-00094-y. DOI <https://doi.org/10.7759/s44389-026-00094-y>