



Using Generative AI to Simulate Patient History-Taking in a Problem-Based Learning Tutorial: A Mixed-Methods Study

Allison Mool¹ · Jacob Schmid¹ · Thomas Johnston¹ · Kodi J. Smith McCoy¹ · Zachary Patterson¹ · Hannah Feldt¹ · William Thomas² · Emma Fenner³ · Kevin Lu¹ · Raya Gandhi¹ · Adam Western¹ · Brendan Seabold¹ · Daniel Vollmer⁴ · Roshan Nallaveetil⁴ · Anthony Fanelli⁴ · Logan Schmillen⁴ · Shelley Tischkau⁵ · Anna T. Cianciolo⁶ · Pinckney Benedict⁷ · Richard Selinfreund⁸

Received: 2 May 2025 / Revised: 16 September 2025 / Accepted: 11 November 2025
© The Author(s) 2025

Abstract

Patient case design in problem-based learning (PBL) must balance realism, educational value, and feasibility. Generative artificial intelligence (genAI)-enabled virtual patients may facilitate this balance, but their use must be examined. In March 2024, 37% of second-year students at a small, community-based medical school in the Midwestern U.S. participated in the pilot implementation of a genAI-enabled virtual patient in a required PBL tutorial. Two groups of students ($N=13$) opened the patient case by interviewing a genAI-enabled avatar. For comparison, two groups ($N=13$) gathered patient history information using the school's legacy system, a searchable multimedia database [Electronic PBL Module (ePBLM)]. All groups had the same expert faculty facilitator. The groups' interactions were examined using descriptive observation. Students' perceptions and their recall of patient history information were assessed using locally developed instruments. GenAI presented essential case content accurately, occasionally deviating from non-essential content. GenAI students spent approximately 10 min longer on history-taking than ePBLM students, collaboratively troubleshooting interaction with the AI. GenAI students rated the clinical accuracy and teamwork aspects of their experience significantly higher than their previous experiences using ePBLMs, however, they treated the avatar like a sophisticated and engaging "question base," rather than a virtual patient. Patient history information recall was near ceiling across conditions. Our study provides clues for supporting clinical learning when representing patients realistically within PBL tutorials. Enhancing students' history-taking from genAI-enabled virtual patients should include framing the experience:

Earlier versions of this work were presented at the 2024 Central Group on Educational Affairs Spring Meeting, Milwaukee, Wisconsin, April 3–5, 2024 and at the SIU-SOM 14th Annual Teaching and Learning Symposium, Springfield, IL, April 11, 2024. An earlier version of this manuscript appears as a preprint in medRxiv—<https://www.medrxiv.org/content/10.1101/2024.05.02.24306753v1> (Last accessed 17 February 2025).

Extended author information available on the last page of the article

prompting students to experiment with AI's capabilities and helping them recognize and adapt to its limitations.

Keywords Problem-based learning · Generative artificial intelligence · History-taking · Virtual patients

1 Introduction

In a medical problem-based learning (PBL) curriculum, patient cases provide the organizing structure, intended to motivate students' intrinsic curiosity and provide a platform for exercising higher-order thinking, including clinical reasoning (Barrows & Tamblyn, 1980; Wee et al., 2001). Patient cases must stimulate the problem-driven collaboration that defines PBL—questioning, explanation and elaboration, identification of knowledge gaps, and discovery of physical mechanisms of action—to achieve its integrated knowledge and clinical learning objectives (Cianciolo & Regehr, 2019). Patient case design is therefore a key concern for PBL educators, and has been investigated for decades (Kim et al., 2006). Research emphasis in case design has evolved as the opportunities and challenges of implementing PBL curricula have, prompted by developments in technology-assisted instruction (e.g., Kamin et al., 2006; Ryan & Koschmann, 1994; Strømsø et al., 2007) and concerns about low student motivation (Dawood et al., 2021; Pyle & Hung, 2019).

Despite the longevity of this research topic, the question of how to represent patient cases to achieve PBL's learning objectives—with possibilities ranging from paper and interactive multimedia cases to live patient actors—remains open. It is widely recognized that quality cases are realistic, ideally based on real patients and set in a medical context (Azer et al., 2012; Cohen et al., 2017; Wee et al., 2001). Realistic patient representation is considered more engaging and provides greater opportunity for practicing clinical reasoning skills, such as history-taking and diagnosis (Azer, 2007; Barrows & Tamblyn, 1980; Cook & Triola, 2009). However, too much realism relative to students' level of training may make the simulated patient encounter overwhelming or distracting, which detracts from the case's knowledge-acquisition objectives (Barrows & Tamblyn, 1980; Bearman, 2003; Cook & Triola, 2009).

At a time when student engagement in PBL is flagging (Dawood et al., 2021; Kilgour et al., 2016; Walling et al., 2017) and technologies for representing patients are burgeoning (Chu & Goodell, 2024; Potter & Jeffries, 2024; Sardesai et al., 2024; Webb, 2023), researching new ways to operationalize realism seems a reasonable step toward modernizing and sustaining PBL for future learners. For instance, research suggests that enhancing the patient's *social presence* (i.e., physical, psychological, and social proximity to learners) may motivate students to discuss the case and engage in independent research (Pyle & Hung, 2019). Manipulating social presence via changes to case modality (written text, audiovisual, face-to-face interaction) and authenticity (real vs. fictional patient) produced compound effects on individual and collaborative activity during a university-level health science PBL tutorial (Pyle & Hung, 2019). Recent advancements in generative artificial intelligence (genAI) may offer another way to manipulate social presence in a PBL tutorial, allowing students to interact with a virtual patient based on a real clinical case, using a modality that can feel psychologically and emotionally immediate (Smith et al., 2025).

Among other capabilities, genAI combines machine learning technologies and large language models to generate realistic interlocutors. Clinical learner interactions with genAI-enabled patients currently range from interactive chat (Potter & Jeffries, 2024; Webb, 2023) to voice-to-voice communication with 3D-animated avatars (Chu & Goodell, 2024; Sardesai et al., 2024). Pilot studies suggest that learners perceive genAI-enabled patients as realistic (Chu & Goodell, 2024; Webb, 2023) and generally rate their usability, accuracy, and utility favorably (Potter & Jeffries, 2024; Sardesai et al., 2024). However, genAI development can be time-consuming and expensive (Chu & Goodell, 2024), and the behavior of this rapidly evolving technology can be unpredictable, inaccurate, and biased (Potter & Jeffries, 2024; Sas, 2024). Bias in genAI representations of minoritized individuals may subvert efforts to promote more patient-centered learning by incorporating diversity into PBL cases (Turbes et al., 2002; MacLeod, 2011; To et al., 2016; Krishnan et al., 2019). In addition, interaction with genAI can facilitate feelings of social and emotional connection that are powerful but not genuine, and possibly harmful to forming real human relationships (Smith et al., 2025).

Various forms of virtual patient—including PBL cases (Barrows & Tamblyn, 1980)—have long been used to develop clinical reasoning skills, particularly history-taking and diagnosis (Cook & Triola, 2009; García-Torres et al., 2024); however, historical forms largely constrained realism to something the educator could control (e.g., high-fidelity, but prefabricated patient responses to learner queries). Using genAI in PBL promotes realistic clinical interactions, but it weakens educator control and introduces new risks of inaccuracies and bias, with unknown impact on students or their learning. Our exploratory study therefore investigated what happened when a genAI-enabled virtual patient was introduced within a PBL tutorial under semi-controlled conditions. We observed interactions among students and the genAI during patient history-taking, and we examined these interactions in the context of students' perceptions of learning quality and recall of case content. We aimed to characterize the impact of using a genAI-enabled patient on PBL learning process, including potential liabilities and associated targets for risk mitigation. Research of this nature is needed as a foundational step to ensuring meaningful and sustainable educational intervention (Hamza & Regehr, 2021).

2 Method

2.1 Overview

This study used a between-groups, mixed-methods approach (Schifferdecker & Reed, 2009) to explore medical student interactions with a genAI-enabled patient during a PBL case opening, using interactions with multimedia patient case materials for comparison. Two student groups conducted voice-to-voice history-taking with a virtual patient avatar (genAI condition). Two others gathered patient history information from a database using keyword search (Electronic PBL Module [ePBLM] condition). We used descriptive observation of audio recordings to characterize students' interactions with both patient modalities. For context, we also surveyed students' perceptions of their learning experience and their recall of patient history information.

2.2 Context/Setting

This study occurred in March 2024 at a small, community-based medical school in the Midwestern U.S., where the 2-year basic science curriculum is delivered using hybrid PBL (Armstrong, 1997). PBL tutor groups comprise 6–8 students and a facilitator who meet multiple times weekly to exercise clinical reasoning with real-world patient cases, discuss the case content and identify knowledge gaps, and develop shared understanding of the patient’s condition and treatment (Barrows, 1994; Cianciolo, et al., 2016; Dawood, et al., 2021; Koschmann, et al., 2012; Ryan & Koschmann, 1994; Walker, et al., 2015). Students collaboratively “open” a case by gathering patient history information, physical exam findings, and test results, forming a differential diagnosis and identifying self-directed learning objectives. Students “close” the case a few days later by discussing their self-directed learning findings, the implications of their new knowledge for understanding the patient’s condition, and the patient’s care outcomes. Our school pioneered ePBLMs 30 years ago (Ryan & Koschmann, 1994), and they are used throughout our PBL curriculum.

2.3 Participants

Participants comprised 26 second-year medical students (37% of the Class of 2026), a convenience sample recruited via word of mouth by their peers on our investigator team (RG, KL, AM, JS). There were no inclusion/exclusion criteria. Among the 14/26 participants who provided demographics, 79% self-identified as women and 20% self-identified as belonging to a minoritized racial/ethnic group.

Participation involved opening the first required case of the final unit in our basic science curriculum (Endocrinology, Reproduction, and Gastroenterology [ERG]), approximately 3 months prior to taking U.S. Medical Licensing Examination, Step 1. Participation offered volunteers an alternative to opening the case with their assigned tutor groups, possibly using a genAI-enabled simulated patient. Volunteers signed up for a time slot convenient for them, with knowledge of who else had selected the same time slot, but without knowledge of the condition (genAI or ePBLM) they would be assigned to. After sign-up, the medical student co-investigators assigned conditions to time slots based on their own availability and preference for proctoring.

This study was deemed non-human subjects research by the Springfield Committee on Research Involving Human Subjects (6 February 2024). The facilitator for all 4 groups, an ERG faculty member on the investigator team—locally considered a master PBL tutor with 16 years’ experience—described the study to participants and received their verbal consent to participate with audio-recording (accomplished via WebEx). Participants in all groups were offered pizza and drinks during their tutorial. No other incentives were provided. Following participation, volunteers reported back to their assigned tutor group to close the case and complete the ERG unit.

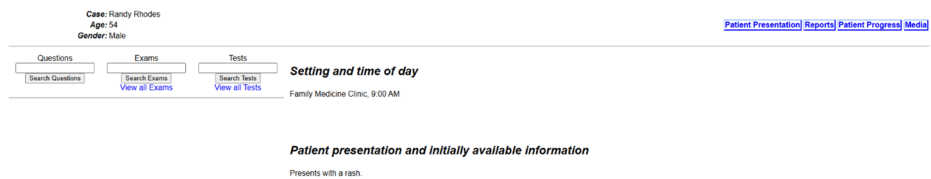


Fig. 1 Screen capture of the Year 2 medical student user interface for the Randy Rhodes electronic problem-based learning module (ePBLM), March 2024

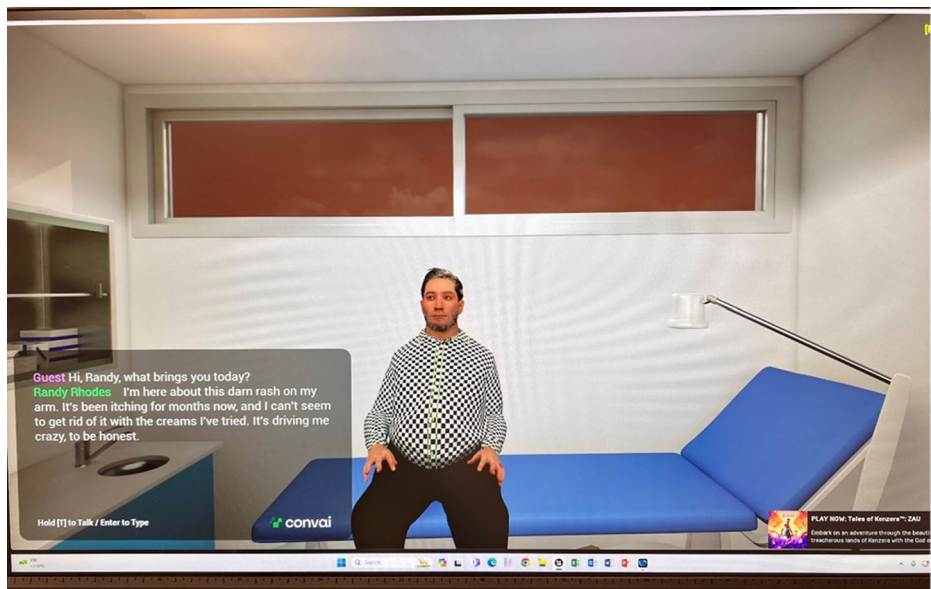


Fig. 2 Screen capture of the Year 2 medical student user interface for the generative AI-enabled patient, Randy Rhodes, a 3D-animated avatar in a clinical exam room developed using Unreal Engine 5.0 and Metahuman, March 2022

2.4 Materials

2.4.1 Patient Case—Randy Rhodes

Randy Rhodes is a 54-year-old man who presents to the family medicine clinic with a rash on his arm. His diagnosis is Diabetes Mellitus Type 2. Based on a real patient seen in our school's clinic, the case is designed to teach pathology, pharmacology, and other topics associated with this diagnosis (e.g., metabolic syndrome) as 1 of 20 cases comprising the ERG PBL curriculum. Participants interacted with Randy Rhodes to gather patient history information either via ePBLM or genAI:

ePBLM. ePBLM cases allow students to gather information serially, using keyword searches to simulate history-taking, physical examination, testing/imaging, and procedures (Fig. 1) (Ryan & Koschmann, 1994). Answers returned to students' queries are presented as

realistically as possible using multimedia. For instance, ePBLMs return scripted responses to patient history queries (e.g., “I’ve had this rash for 3 months.”).

GenAI. Students gather patient history information from the genAI-enabled patient via voice-to-voice interaction with a 3D-animated avatar situated in a simulated clinical exam room and developed using Unreal Engine 5.0 and Metahuman (Fig. 2). The avatar’s technical infrastructure comprises a combination of commercial platforms. A local speech-to-text model transcribes students’ spoken queries, and we use ElevenLabs APIs to synthesize the avatar’s custom voice responses. Simultaneously, the D-ID facial animation platform renders the avatar’s visual expressions in real-time to create a full audiovisual experience. To produce the avatar’s responses, we use API calls to OpenAI’s GPT-4o large language model, which processes students’ transcribed queries based on custom agent instructions informed by the faculty-generated case materials contained in the Randy Rhodes ePBLM database (see Online Resource 1 for details). The scope of the avatar’s responses is shaped by “guardrails” placed to optimize the case’s educational value. For instance, one guardrail prevents Randy from revealing his diagnosis directly to students, while others ensure accurate presentation of pertinent positive findings. Within these limits, the genAI is allowed to respond adaptively to students’ questions, as overly restricting large language models produces robotic, unrealistic interlocutors (Horkan, 2024; Tancredi, 2024).

2.4.2 Descriptive Observation Notes

Once all 4 tutor groups had met, 5 co-investigators conducted descriptive observation (Spradley, 1980) of the audio-recordings, documenting their notes using a locally created template (Online Resource 2). The template prompted observers to (1) document a reflexivity reflection; (2) describe noteworthy occurrences observed in the audio-recording; and (3) report their reactions to these occurrences. We designed the template this way to allow observers to voice their opinions and reflections, but in a bracketed way that preserves objectivity in how noteworthy occurrences are documented. The lead observer (AC) was a behavioral scientist with experience conducting observational research of PBL tutorials (Cianciolo et al., 2016) who was uninvolved in the development of the genAI-enabled patient and inclined to approach genAI with caution. The other observers (HF, AM, ZP, KS) were medical student co-investigators who were involved in the genAI-enabled patient’s development and had insight into how it was supposed to behave. Collectively, this team was well-positioned to understand student behavior and to identify opportunities and limitations in patient modality across groups.

2.4.3 Learner Perceptions Survey

This 8-item survey assessed participants’ perceptions of PBL tutorials using either ePBLM or genAI. Participants used a scale of 1 to 5 to rate how positively they view several aspects of tutorial quality and impact (impact on improving clinical knowledge, clinical simulation accuracy, impact on clinical knowledge retention, enjoyability, stressfulness, allowance for teamwork, interestingness, impact on confidence) when either patient modality is used. This instrument was locally developed due to the lack of an existing instrument relevant to our curriculum. However, the general dimensions of perception assessed (accuracy/realism, utility/perceived impact, engagement) were judged by the behavioral scientist on our team

(AC) to be typical of those examined in other studies of simulated patients (e.g., García-Torres et al., 2024; Sardesai et al., 2024; Wimmer et al., 2014). The composite score for each form comprised the sum of ratings across all 8 items (item 5 was reverse-coded), ranging from 8 to 40 and representing how much positive perception participants have of their overall PBL tutorial experience using each patient modality (Online Resources 3–4).

2.4.4 Patient History Information Quiz

This 11-item, short-answer quiz (Online Resource 5) assessed knowledge of Randy Rhodes's demographic information, medical history, and symptoms immediately following the PBL tutorial and 2 weeks later. We assessed knowledge this way to explore whether interacting with genAI can make patient case information more memorable, for longer, potentially serving students later in their training the way a prior (real) patient experience might. Lacking an existing instrument to assess case-specific knowledge, AM developed the items locally using the faculty-developed case materials. Item quality (i.e., consistency with the case, readability) was verified by two members of the research team with extensive Randy Rhodes case experience (ST) and performance assessment research experience (AC). Quizzes were graded by 2 student co-investigators (AM, JS) masked to the condition (ePBLM or genAI) and timing (immediate or delayed) when the quiz was completed. According to a faculty-verified key, each short answer could receive 0 points, ½ point, or 1 point, and the possible range of total scores was 0 to 11, representing the quantity of patient history information recalled from the Randy Rhodes case. Both co-investigators independently graded the first 5 quizzes to establish consistency with the rating criteria, after which 1 co-investigator (JS) graded the remaining 52 quizzes.

2.5 Procedure

All 4 groups met in a room arranged to mimic a typical PBL classroom and equipped with the tools needed to interact with the simulated patient: a keyboard and mouse to search the ePBLM database and a push-to-talk microphone to interview the AI. After consent, participants completed the Learner Perceptions Survey-ePBLM, rating their perceptions of ePBLM tutorials generally based on their nearly 2 years' experience in our PBL curriculum. Surveys were proctored by 1 of 4 student co-investigators (RG, KL, AM, JS) who also answered participant questions, operated the patient-interaction tools (i.e., push-to-talk microphone, the keyboard and mouse), and managed audio-recording. With minimal direction from the faculty facilitator, all 4 groups conducted the case opening. Immediately following, participants completed the Patient History Information Quiz and retook the Learner Perceptions Survey, rating their perceptions of PBL tutorials using the type of simulated patient they had just interacted with (ePBLM or genAI). Two weeks later, participants individually retook the Patient History Information Quiz, proctored by a student co-investigator (AM, KL, or RG), at times and campus locations of convenience.

2.6 Data Analysis

2.6.1 Observation Data

Observation notes were analyzed inductively by 4 members of the 5-person observation team (AC, HF, ZP, KS). For each group, the lead analyst (AC) compiled the independent observers' notes into one set, arranging these notes chronologically to create detailed, multi-perspective documentation of noteworthy events and associated bracketed reflections that occurred during each tutorial. The noteworthy events documented overlapped substantially among the researchers. Next, all analysts independently reviewed each set of compiled field notes to identify regularities in the interactions among people and case materials. Coded as regularities were recurring interactional behaviors such as consistencies in patient interview strategies or commonalities among occurrences of paused activity or silence. Analysts also coded interactional behaviors potentially reflective of how students oriented to the patient case or each other (e.g., referring to the patient with personal pronouns or seeking input from peers). All analysts worked collaboratively to develop codes iteratively through a process of independent coding, consensus-building discussion, codebook refinement, and re-coding until a consensus coding for all four groups was achieved (Online Resource 6). Finally, all analysts collaboratively reviewed the codes within and across participant groups to describe patterns in PBL learning process associated with type of patient representation.

2.6.2 Survey/Quiz Data

To examine the effect of condition (ePBLM vs. genAI) on participants' perceptions and recall of patient history information, we fitted linear regression models using ordinary least squares (OLS) estimation. All statistical tests were two-tailed with $\alpha=0.05$. Confidence intervals were calculated using a Wald t-distribution approximation. All analyses were performed using R version 4.5.0 using standard statistical and visualization packages (lme4, emmeans, sjPlot, gtsummary, ggplot2, and related packages).

3 Results

3.1 Tutorial Time

GenAI tutorials took longer (55–65 min) than ePBLM tutorials (36–39 min), in part due to time spent history-taking. GenAI groups spent approximately 10 min longer taking the patient history than ePBLM groups (24–30 min vs. 16–19 min for genAI and ePBLM groups, respectively). In addition, only genAI groups concluded the tutorial by communicating the group's diagnosis and initial management plans to the patient (10–12 min).

3.2 Student History-Taking: Interactions with the Patient Representation

ePBLM students collaboratively identified what patient history information to gather (e.g., whether the rash was painful) and nominated keywords most likely to produce it (e.g., "Pain."). Occasionally, difficulty identifying the right keyword frustrated their efforts, or

the sought-after information was incomplete or discovered when looking for something else (e.g., “Students troubleshoot to try and find what dangerous substances Randy may be exposed to. They try “exposures,” then “hazards,” but find the listing under “occupation.”). All observers noted that the flow of history-taking in these groups was typical for tutorials in our PBL curriculum.

GenAI students collaboratively questioned the AI and appeared to find the avatar’s appearance, voice quality, and behavior plausible; no language or interactions suggested an aversive reaction to the patient modality overall. Most of the students’ questions were closed-ended, seeking a yes/no, single-word, or short-answer response (e.g., “Has [the rash] been getting worse over the past 3 months?”), and they often used jargon or formal, medical tone. In general, students seemed more focused on how to gather history information efficiently than on approximating interaction with a real human. However, Group 1 adopted a more open and experimental stance than Group 2, more often using the patient’s name in their interview questions (e.g., “Randy, does anything make the rash better?”), inferring sentence or agency from the AI’s behavior (e.g., “He’s pulling an attitude.”), and reacting to the AI’s responses appreciatively, empathetically, or playfully. Uncertainty regarding how realistically to engage clinically with the AI was illustrated by a Group 1 student’s question to the tutor: “Can we give him advice or is it [the AI] just, like, a question base?” Both groups encountered some difficulty with voice-to-voice communications, causing frustration (e.g., temporary withdrawal from asking questions), but observers did not note persistent effects of these technical challenges in either group.

3.3 GenAI-Enabled Patient—Content Representation

The genAI-enabled patient presented pertinent positive findings consistently with the ePBLM. Deviations occurred when the AI responded to questions for which the ePBLM content was minimal or provided a non-pertinent negative finding. For example, the ePBLM content returned if students queried about headaches was “Only on occasion.” When genAI students asked about headaches, it replied: “I haven’t had any significant headaches to note” in one group, and “[headaches] for a few months now, on and off” in the other. When queried regarding drug and alcohol use, the ePBLM returned “Rarely a beer or two on the weekend.” By contrast, the genAI reported occasional recreational marijuana use.

Although Group 1 was observed as more engaged in interacting with the AI, it did not appear to respond in kind. The AI tended to provide more forthcoming responses to Group 2’s questions, and these responses sometimes provided pertinent positive findings incidentally (e.g., “AI also volunteers an increase in urinary frequency in response to this question [about fatigue].”). The AI in Group 2 also tended to embellish its responses more often, which seemed to feature rapport-building strategies (e.g., “Randy describes the rash on his arm, finishing with: ‘It’s driving me nuts, man.’”). By contrast, the AI’s responses in Group 1 tended to use language that paralleled the students’ questions, and sometimes delivered these responses in an irritable tone of voice, which the students found surprising, but non-disruptive to the group dynamic.

Table 1 Linear regression model of survey data reflecting year 2 medical students' perceptions of simulated patients in Problem-Based Learning—Estimated marginal Means, March 2024

Time	Condition	Mean	SE	95% CI	t(48)
Pre	ePBLM	29.77	1.12	[27.52, 32.02]	26.58
Pre	genAI	28.38	1.12	[26.13, 30.64]	25.35
Post	ePBLM	30.00	1.12	[27.75, 32.25]	26.79
Post	genAI	34.38	1.12	[32.13, 36.64]	30.71

ePBLM, Electronic Problem-Based Learning Module; genAI, Generative Artificial Intelligence

Table 2 Linear regression model of survey data reflecting year 2 medical students' perceptions of simulated patients in Problem-Based Learning—Pairwise Contrasts, March 2024

Level1	Level2	Diff	SE	95% CI	t(48)	p
Pre ePBLM	Pre genAI	1.38	1.58	[− 1.80, 4.57]	0.87	0.464
Pre ePBLM	Post ePBLM	− 0.23	1.58	[− 3.41, 2.95]	− 0.15	0.885
Pre genAI	Post ePBLM	− 1.62	1.58	[− 4.80, 1.57]	− 1.02	0.464
Pre ePBLM	Post genAI	− 4.62	1.58	[− 7.80, − 1.43]	− 2.91	0.016
Pre genAI	Post genAI	− 6.00	1.58	[− 9.18, − 2.82]	− 3.79	0.003
Post ePBLM	Post genAI	− 4.38	1.58	[− 7.57, − 1.20]	− 2.77	0.016

p value adjustment method: Benjamini and Hochberg (1995)

ePBLM, Electronic Problem-Based Learning Module; genAI, Generative Artificial Intelligence

3.4 Student History-Taking: Group Interaction

Observers noted a high level of interaction among students, regardless of condition. ePBLM students contributed keywords, asked each other clarification questions regarding the information gathered, and conferred on its implications. GenAI students had to phrase their questions naturalistically, and they worked collaboratively on phrasing and troubleshooting incomplete or unexpected responses. GenAI Group 1 showed a much larger degree of collaborative problem solving than Group 2, both to communicate with the virtual patient and to interpret case content. No obvious differences were observed in tutor interjections across conditions, though there were fewer in the most interactive group (i.e., genAI Group 1).

3.5 Learner Perceptions Survey Data

Linear regression modeling of the learner perceptions survey data (formula: survey_total~time * condition) revealed a significant effect of time (pre- vs. post-tutorial) for the genAI condition only [$\beta=5.77$, 95% CI [1.27, 10.27], $t(48)=2.58$, $p=.013$; Std. $\beta=1.27$, 95% CI [0.28, 2.27, see also Tables 1 and 2). That is, genAI students rated their present PBL learning experience significantly higher overall than they rated their previous PBL learning experiences using ePBLMs, especially with respect to authenticity (Table 3). On average, students who gathered patient history information from the genAI-enabled patient rated the clinical accuracy of their present PBL tutorial 1.6 points higher than that of their previous tutorials using ePBLMs. The regression model explained a statistically significant proportion of the variance in survey total scores [$R^2=0.25$, $F(3, 48)=5.39$, $p=.003$, adj. $R^2=0.21$].

Table 3 Item-Level data from survey assessing year 2 medical students’ perceptions of simulated patients in Problem-Based Learning—Mean (Standard Deviation), March 2024

Survey Item	Pre-Tutorial (N=26)		Post-Tutorial (N=26)	
	genAI (N=13)	ePBLM (N=13)	genAI (N=13)	ePBLM (N=13)
Improves clinical knowledge	3.8 (0.6)	3.8 (0.9)	4.2 (0.6)	3.6 (0.7)
Simulates clinical experiences accurately	2.6 (1.0)	3.1 (1.0)	4.2 (0.6)	3.5 (0.9)
Helps clinical knowledge retention	3.6 (1.0)	3.6 (1.0)	4.3 (0.6)	3.5 (0.9)
Is enjoyable	3.5 (0.8)	3.7 (0.6)	4.7 (0.5)	3.8 (0.8)
Is stressful*	4.1 (0.5)	3.9 (0.5)	3.8 (0.8)	4.2 (0.6)
Allows for teamwork	3.8 (0.4)	4.0 (0.4)	4.5 (0.7)	4.0 (0.6)
Stimulates interest in learning the disease at hand	3.4 (0.9)	3.8 (0.7)	4.4 (0.5)	3.8 (0.8)
Improves confidence interacting with this chief complaint	3.5 (1.0)	3.8 (0.7)	4.2 (0.8)	3.5 (0.7)
Total	28.4 (0.4)	29.8 (0.3)	34.4 (0.3)	30.0 (0.3)

*This item is reverse-coded such that a higher score represents lower perceived stress
genAI, Generative Artificial Intelligence; ePBLM, Electronic Problem-Based Learning Module

Table 4 Linear regression model of data from a quiz assessing year 2 medical students’ knowledge of patient case history information (Immediate and delayed Recall), March 2024

Predictors	Estimates	Std. Error	CI	t	p
(Intercept)	9.40	0.30	8.80 – 10.01	31.24	<0.001
Time: delayed	-1.46	0.35	-2.16 – -0.76	-4.20	<0.001
Condition: genAI	0.69	0.35	-0.01 – 1.39	1.99	0.052

Observations 52

R² / R² adjusted 0.306/0.278

F(2, 49)=10.82, p<.001

genAI, Generative Artificial Intelligence, Bolded values are significant at p < 0.001

3.6 Patient History Quiz Data

Immediate recall performance in both conditions was near ceiling (9.40 and 10.10 out of 11 possible points in ePBLM and genAI, respectively), and delayed recall was also very high (7.94 and 8.63 in ePBLM and genAI, respectively). Linear regression modeling, with time (immediate vs. delayed) and condition (ePBLM vs. AI) used as predictors (formula: quiz~time+condition), explained a statistically significant proportion of the variance in scores [R² = 0.31, F(2, 49)=10.82, p<.001, adjusted R² = 0.28]. A significant effect was found for time only (Table 4). Across conditions, students’ recall decreased significantly from immediate to delayed assessment, and the rate of forgetting did not differ by condition. Observations 52R² / R² adjusted 0.306 / 0.278F(2, 49)= 10.82, p < .001 Abbreviations: genAI, Generative Artificial Intelligence

4 Discussion

This study followed the tradition of directly observing learner interactions with virtual patients (see, e.g., García-Torres et al., 2024), and it is among the few that have directly observed the impact of manipulating patient modality in PBL (Kamin et al., 2006; Pyle & Hung, 2019). Our findings suggest that a genAI-enabled patient can accurately represent essential elements of PBL case content (i.e., pertinent positive findings), stimulate the problem-driven interactions definitive of PBL, and allow experimentation with interview questioning. Students in the genAI condition spent more time on history-taking, troubleshooting how to pose questions and collaboratively interpret the AI's responses. Students who interacted with the genAI-enabled patient reported that their learning experience was more enjoyable and facilitative of teamwork than did the students who interacted with multimedia case content (i.e., ePBLM). However, students' interactions with the genAI fell short of realistic patient interviewing; their questioning strategies generally emphasized efficiency and suggested perception of the genAI-enabled patient as a sophisticated and engaging "question base."

There are several possible reasons why our participants did not interview the genAI-enabled patient like a real human, even as some students seemed to anthropomorphize it. First, voice communication with the AI was mediated by a student co-investigator, who operated the push-to-talk microphone, which potentially stifled interaction with the avatar and reduced its social presence with the students. Technical difficulty with voice-to-voice communication also disrupted the flow of history-taking and sometimes caused frustration. Second, as rising third-year students, our participants were experienced with querying ePBLMs, and likely brought this framework to interviewing the genAI-enabled patient (Bateman et al., 2013). We likely reinforced this frame by incorporating our study into required PBL curriculum. Finally, to maintain the study's exploratory approach and avoid influencing students' attitudes towards the virtual patient, the faculty facilitator purposefully did not offer students a divergent perspective on how to interact with the AI.

Studies exploring the impact of realistic patient representation in PBL have used high-fidelity simulation in auxiliary fashion, either as a stimulating introduction to collaborative problem-solving or as a practice opportunity afterward (Euliano & Mahla, 1999; Nash & Crowther, 2018; Wimmer et al., 2014). Our study provides clues for supporting clinical learning when representing patients realistically *within* PBL tutorials. Our participants' history-taking comprised predominantly closed-ended questioning and was not reflective of ideal patient-centered communication (Han et al., 2022; Koschmann et al., 2019). We are improving our voice-to-voice communication system to allow more direct, naturalistic interactions, and our Randy Rhodes avatar now looks more realistic. However, students may need encouragement from tutors to engage with the genAI-enabled patient's capabilities to act humanlike (Bateman et al., 2013). This encouragement could be reinforced by a broad orientation to genAI, highlighting opportunities for its use in PBL, which aligns with recommendations for teaching medical students about AI (Preiksaitis & Rose, 2023). Finally, students may be more inclined toward humanistic interactions with genAI-enabled patients if implementation begins early in the curriculum, perhaps in the first year, before an alternative framework can take hold (Bateman et al., 2013).

When representing patients in PBL, realism, clinical accuracy, and resource efficiency are chief concerns (Azer, 2007; Azer et al., 2012; Bokken et al., 2010; Cohen et al., 2017;

Cook & Triola, 2009; Wee et al., 2001). If framed and facilitated effectively, genAI-enabled patients could potentially achieve greater social presence with students, but operationalizing realism in this way comes with the risks of unpredictability, inaccuracy, bias, and inappropriate emotional dependence (Sas, 2024; Potter & Jeffries, 2024; Smith et al., 2025). However, too much control obviates genAI's capacity to act humanlike (Horkan, 2024; Tancredi, 2024). Using genAI effectively in PBL therefore raises an interesting philosophical question (Schrage & Kiron, 2025): how much and what kind of deviation from faculty control disrupts PBL's goals of developing knowledge, clinical skill, teamwork, and self-directed learning (Barrows & Tamblyn, 1980)? In this study, the genAI embellished its responses to questions with minimal content in the ePBLM, usually non-pertinent negative findings. This embellishment sometimes produced a content deviation flagged by observers and even the tutor, who was very familiar with the case. Although interacting with the genAI-enabled patient stimulated problem-focused collaborative interactions typical of PBL (Cianciolo & Regehr, 2019), the impact of content deviation should be examined systematically across different types of cases, particularly those with more pertinent negative findings. Finally, the inevitability of content deviation when using genAI suggests a new tutor role in genAI-enhanced PBL: helping students stay on track when content deviation tempts distraction.

4.1 Limitations and Future Directions

This study involved one case at one medical school, and all groups were facilitated by a single, experienced faculty member enthusiastic about using genAI in PBL. Although students were randomly allocated to condition, they all volunteered to participate in a study that offered the chance to use genAI instead of the ePBLM; one could reasonably assume our participants also were enthusiastic about using genAI in PBL. Moreover, our medical school has a longstanding and well-documented history of research and innovation in PBL, including PBLM (Barrows, 1994; Cianciolo et al., 2016; Dawood et al., 2021; Koschmann et al., 2012; Ryan & Koschmann, 1994; Walker et al., 2015). We therefore conducted our study under "best case" conditions that may be unique to a particular group of people at our school. In addition, we used locally developed instruments to assess knowledge and perceptions of learning experience, which precludes generalization about our participants' outcomes beyond our local setting. Although more than one-third of the class participated in our study, our short-term assessment of knowledge retention and the numerous sources of diversity among PBL groups, including tutor behavior and the group's prior knowledge (Cianciolo, et al., 2016; Schmidt & Moust, 2000), further limit generalizability even within our own setting. In addition, the small number of groups in our study precluded statistical control of group effects on the patient history quiz and perceptions survey data. Given our sample, we selected a simpler analytical approach to optimize statistical and explanatory power.

However, as first adopters of genAI-enabled patients in PBL, our aim with this study was not generalizability. Rather, we sought to lay the groundwork for systematically observing the adoption of genAI-enabled patients in PBL to facilitate their meaningful and lasting use for education (Hamza & Regehr, 2021). Next steps for our research involve diversifying the conditions under which genAI is used in our PBL curriculum and using video data to enrich our observational analysis. Our observations in the present study relied on audio-recorded data only, but video data would allow more comprehensive observations of how students

interact with the patient representation and with each other. In addition, assessments that capture differences in learning issue identification (Dolmans et al., 1993), critical thinking (Kamin et al., 2001) and clinical reasoning (Cook & Triola, 2009) when students interact with a genAI-enabled virtual patient vs. a multimedia database would allow more meaningful and consistent insight into the conduct of modern PBL. Focused investigation of how genAI-enabled patients in PBL affect students' clinical information gathering and use would integrate the virtual patient literature (Cook & Triola, 2009; García-Torres et al., 2024) with the literature on PBL case design (Azer, 2007; Azer et al., 2012; Bokken et al., 2010; Cohen et al., 2017; Pyle & Hung, 2019; Wee et al., 2001) to inform nuanced, adaptive pedagogical practices that optimize technology use in meeting PBL's learning objectives.

5 Conclusion

To our knowledge, this study is the first to examine what happened when a genAI-enabled virtual patient was implemented in a PBL tutorial. Our study advances understanding of patient representation in PBL by exploring how an emerging patient simulation modality affects learner interactivity and engagement with case content. Our study also adds to research exploring the impact of genAI on education by directly observing learning process instead of relying on activity data and/or self-reported learning experiences. This work represents the first step in a program of research aimed at leveraging the unique capabilities of genAI to engage medical learners and promote knowledge and skill acquisition. We also seek to deepen conceptual understanding of patient case design and interactional processes in PBL, using genAI as a platform to investigate the contextual factors that influence the nature and extent of learner engagement.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10758-025-09929-4>.

Acknowledgements The authors would like to thank Don Torrey, PhD, for his guidance and continued focus related to research strategy; Debra L. Klamen, MD, MHPE, for her continued support for this project; and Sean McGinity BS, BA for his support to the quantitative data analysis. Finally, the authors would like to acknowledge the medical students from the Class of 2026 who participated in this study.

Author Contributions Allison Mool: investigation, methodology, formal analysis, writing original draft, writing—review & editing. Jacob Schmid: conceptualization, investigation, methodology, writing—review & editing. Thomas Johnston: methodology, writing original draft, writing—review & editing. Kodi J. Smith McCoy: formal analysis, methodology, software, writing original draft, writing—review & editing. Zachary Patterson: formal analysis, methodology, software, writing original draft, writing—review & editing. Hannah Feldt: formal analysis, methodology, software, writing original draft, writing—review & editing. William Thomas: formal analysis, investigation, methodology, writing—review & editing. Emma Fenner: conceptualization, writing—review & editing. Kevin Lu: investigation, resources, writing—review & editing. Raya Gandhi: formal analysis, investigation, resources, writing—review & editing. Adam Western: conceptualization, resources, writing—review & editing. Brendan Seabold: resources, writing—review & editing. Daniel Vollmer: software, writing—review & editing. Roshan Nallaveetil: software, writing—review & editing. Anthony Fanelli: software, writing—review & editing. Logan Schmillen: software, writing—review & editing. Shelley Tischkau: conceptualization, investigation, writing—review & editing. Anna T. Cianciolo: conceptualization, formal analysis, writing original draft, writing—review & editing. Pinckney Benedict: conceptualization, methodology, resources, visualization, writing—review & editing. Richard Selinfreund: funding acquisition, methodology, project administration, resources, supervision, writing—review & editing.

Funding This project received funding from 3 internal sources at Southern Illinois University School of Medicine (SIU-SOM): Dean's Start-up Funds; SIU-SOM Concept Development Award; and an SIU-SOM Foundation Grant for Diabetes Research.

Declarations

Competing interests The technology examined in this study is being developed with commercialization as a goal. To prevent the intrusion of commercial bias, the study conduct and results were presented at multiple stages to curriculum committees and at academic conferences.

Ethical Approval This project was deemed non-human subjects research and therefore exempt from oversight by the Springfield Committee on Research Involving Human Subjects (February 6, 2024).

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

- Armstrong, E. G. (1997). A hybrid model of problem-based learning. In D. Boud, & G. Feletti (Eds.), *The challenge of problem-based learning* (2nd ed., pp. 137–150). Kogan Page.
- Azer, S. A. (2007). Twelve tips for creating trigger images for problem-based learning cases. *Medical Teacher*, 29, 93–97. <https://doi.org/10.1080/01421590701291444>
- Azer, S. A., Peterson, R., Guerrero, A. P. S., & Edgren, G. (2012). Twelve tips for constructing problem-based learning cases. *Medical Teacher*, 34, 361–367. <https://doi.org/10.3109/0142159X.2011.613500>
- Barrows, H. S. (1994). Practice-based learning: Problem-based learning applied to medical education. Southern Illinois University, School of Medicine, PO Box 19230, Springfield, IL 62794-9230.
- Barrows, H. S., & Tamblyn, R. M. (1980). *Problem-based learning: An approach to medical education*. Springer.
- Bateman, J., Allen, M., Samani, D., Kidd, J., & Davies, D. (2013). Virtual patient design: Exploring what works and why: A grounded theory study. *Medical Education*, 47(6), 595–606. <https://doi.org/10.1111/medu.12151>
- Bearman, M. (2003). Is virtual the same as real? Medical students' experiences of a virtual patient. *Academic Medicine*, 78(5), 538–545.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1), 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Bokken, L., Rethans, J. J., Jöbsis, Q., Duvivier, R., Scherpbier, A., & van der Vleuten, C. (2010). Instructiveness of real patients and simulated patients in undergraduate medical education: A randomized experiment. *Academic Medicine*, 85(1), 148–154. <https://doi.org/10.1097/ACM.0b013e3181c48130>
- Chu, S. N., & Goodell, A. J. (2024). Synthetic patients: Simulating difficult conversations with multimodal generative AI for medical education. [arXiv:2405.19941](https://arxiv.org/abs/2405.19941). 2024 May 30.
- Cianciolo, A. T., & Regehr, G. (2019). Learning theory and educational intervention: Producing meaningful evidence of impact through layered analysis. *Academic Medicine*, 94(6), 789–794. <https://doi.org/10.1097/ACM.00000000000002591>
- Cianciolo, A. T., Kidd, B., & Murray, S. (2016). Observational analysis of near-peer and faculty tutoring in problem-based learning groups. *Medical Education*, 50(7), 757–767. <https://doi.org/10.1111/medu.12969>

- Cohen, D. A., Newman, L. R., & Fishman, L. N. (2017). Twelve tips on writing a discussion case that facilitates teaching and engages learners. *Medical Teacher*, 39(2), 147–152. <https://doi.org/10.1080/0142159X.2017.1266315>.
- Cook, D. A., & Triola, M. M. (2009). Virtual patients: A critical literature review and proposed next steps. *Medical Education*, 43(4), 303–311. <https://doi.org/10.1111/j.1365-2923.2008.03286.x>
- Dawood, O., Rea, J., Decker, N., Kelley, T., & Cianciolo, A. T. (2021). Problem-based learning about problem-based learning: Lessons learned from a student-led initiative to improve tutor group interaction. *Medical Science Educator*, 31, 395–399.
- Dolmans, D. H., Gijsselaers, W. H., Schmidt, H. G., & Van Der Meer, S. B. (1993). Problem effectiveness in a course using problem-based learning. *Academic Medicine*, 68(3), 207–213.
- Euliano, T. Y., & Mahla, M. E. (1999). Problem-based learning in residency education: A novel implementation using a simulator. *Journal of Clinical Monitoring and Computing*, 15, 227–232.
- García-Torres, D., Ripoll, M. A. V., Peris, C. F., & Solves, J. J. M. (2024). Enhancing clinical reasoning with virtual patients: A hybrid systematic review combining human reviewers and ChatGPT. *Healthcare*, 12(22), 2241. <https://doi.org/10.3390/healthcare1222241>
- Hamza, D. M., & Regehr, G. (2021). Eco-normalization: Evaluating the longevity of an innovation in context. *Academic Medicine*, 96(11S), 48–53. <https://doi.org/10.1097/ACM.0000000000004318>
- Han, H., Hingle, S. T., Koschmann, T., Papireddy, M. R., & Ferguson, J. (2022). Analyzing expert criteria for authentic resident communication skills. *Teaching and Learning in Medicine*, 34(1), 33–42. <https://doi.org/10.1080/10401334.2021.1977134>
- Horkan, W. (2024). Why AI responses vary: Understanding the subjectivity and variability in AI language models. July 23, 2024. Available at <https://horkan.com/2024/07/23/why-ai-responses-vary-understanding-the-subjectivity-and-variability-in-ai-language-models>. Last Accessed on 10 March 2025.
- Kamin, C., O'Sullivan, P. S., Younger, M., & Deterding, R. (2001). Measuring critical thinking in problem-based learning discourse. *Teaching and Learning in Medicine*, 13(1), 27–35. https://doi.org/10.1207/S15328015TLM1301_6
- Kamin, C. S., O'Sullivan, P., Deterding, R. R., Younger, M., & Wade, T. (2006). A case study of teaching presence in virtual problem-based learning groups. *Medical Teacher*, 28(5), 425–428. <https://doi.org/10.1080/014215906006028241>
- Kilgour, J. M., Grundy, L., & Monrouxe, L. V. (2016). A rapid review of the factors affecting healthcare students' satisfaction with small-group, active learning methods. *Teaching and Learning in Medicine*, 28(1), 15–25. <https://doi.org/10.1080/10401334.2015.1107484>
- Kim, S., Phillips, W. R., Pinsky, L., Brock, D., Phillips, K., & Keary, J. (2006). A conceptual framework for developing teaching cases: A review and synthesis of the literature across disciplines. *Medical Education*, 40(9), 867–876. <https://doi.org/10.1111/j.1365-2929.2006.02544.x>
- Koschmann, T., Frankel, R., & Albers, J. (2019). A Tale of two inquiries (or, doing being competent in a clinical skills exam). *Teaching and Learning in Medicine*, 31(3), 258–269. <https://doi.org/10.1080/10401334.2018.1530597>
- Koschmann, T., Kelson, A. C., Feltovich, P. J., & Barrows, H. S. (2012, October). Computer-supported problem-based learning: A principled approach to the use of computers in collaborative learning. In *Cscl* (pp. 83–124). Routledge.
- Krishnan, A., Rabinowitz, M., Ziminsky, A., Scott, S. M., & Chretien, K. C. (2019). Addressing race, culture, and structural inequality in medical education: A guide for revising teaching cases. *Academic Medicine*, 94(4), 550–555. <https://doi.org/10.1097/ACM.0000000000002589>
- MacLeod, A. (2011). Six ways problem-based learning cases can sabotage patient-centered medical education. *Academic Medicine*, 86, 818–825. <https://doi.org/10.1097/ACM.0b013e31821db670>
- Nash, E., & Crowther, J. (2018). Enhancing student learning by combining problem based learning and human patient simulation. *MedEdPublish*, 7, 263.
- Potter, L., & Jefferies, C. (2024). Enhancing communication and clinical reasoning in medical education: Building virtual patients with generative AI. *Future Healthcare Journal*, 11, 100043. <https://doi.org/10.1016/j.fhj.2024.100043>
- Preiksaitis, C., & Rose, C. (2023). Opportunities, challenges, and future directions of generative artificial intelligence in medical education: Scoping review. *Journal of Medical Internet Research - Medical Education*, 9, e48785.
- Pyle, E., & Hung, W. (2019). The role of subject presence type on student motivation in a PBL learning environment. *Advances in Health Sciences Education*, 24(4), 643–663. <https://doi.org/10.1007/s10459-019-09889-2>
- Ryan, C., & Koschmann, T. (1994). The collaborative learning laboratory: A technology-enriched environment to support problem-based learning.

- Sardesai, N., Russo, P., Martin, J., & Sardesai, A. (2024). Utilizing generative conversational artificial intelligence to create simulated patient encounters: A pilot study for anaesthesia training. *Postgraduate Medical Journal*, 100(1182), 237–241. <https://doi.org/10.1093/postmj/qgad137>
- Sas, M. (2024). Unleashing generative non-player characters in video games: An AI Act perspective. In 2024 *IEEE gaming, entertainment, and media conference (GEM)* (pp. 1–4). IEEE. <https://doi.org/10.1109/GEM61861.2024.10585442>
- Schiffedercker, K. E., & Reed, V. A. (2009). Using mixed methods research in medical education: Basic guidelines for researchers. *Medical Education*, 43(7), 637–644. <https://doi.org/10.1111/j.1365-2923.2009.03386.x>
- Schmidt, H. G., & Moust, J. H. C. (2000). Factors affecting small-group tutorial learning: A review of research. In D. H. Evensen, & C. E. Hmelo (Eds.), *Problem-based learning: A research perspective on learning interactions* (pp. 19–51). Lawrence Erlbaum Associates.
- Schrage, M., & Kiron, D. (2025). Philosophy Eats AI. *MIT Sloan Management Review* (Online), 1–7. Available at <https://sloanreview.mit.edu/article/philosophy-eats-ai/>. Last Accessed on 10 March 2025.
- Smith, M. G., Bradbury, T. N., & Karney, B. R. (2025). Can generative AI chatbots emulate human connection? A relationship science perspective. *Perspectives on Psychological Science*, 17456916251351306. <https://doi.org/10.1177/17456916251351306>
- Spradley, J. P. (1980). *Participant observation*. Thomson Learning.
- Strømso, H. I., Grøttum, P., & Lycke, K. H. (2007). Content and processes in problem-based learning: A comparison of computer-mediated and face-to-face communication. *Computer Assisted Learning*, 23(3), 271–282. <https://doi.org/10.1111/j.1365-2729.2007.00221.x>
- Tancredi, L. (2024). AI strategies series: inconsistent and outdated responses. February 21, Available at <https://insight.factset.com/ai-strategies-series-inconsistent-and-outdated-responses>. Last Accessed on 10 March 2025.
- To, M. J., MacLeod, A., & Hwang, S. W. (2016). Homelessness in the medical curriculum: An analysis of case-based learning content from one Canadian medical school. *Teaching and Learning in Medicine*, 28(1), 35–40. <https://doi.org/10.1080/10401334.2015.1108198>
- Turbes, S., Krebs, E., & Axtell, S. (2002). The hidden curriculum in multicultural medical education: The role of case examples. *Academic Medicine*, 77(3), 209–216.
- Walker, A., Leary, H., & Hmelo-Silver, C. (Eds.). (2015). *Essential readings in problem-based learning: Exploring and extending the legacy of Howard S. Barrows*. Purdue University Press.
- Walling, A., Istas, K., Bonaminio, G. A., Paolo, A. M., Fontes, J. D., Davis, N., & Berardo, A. (2017). Medical student perspectives of active learning: A focus group study. *Teaching and Learning in Medicine*, 29(2), 173–180. <https://doi.org/10.1080/10401334.2016.1247708>
- Webb, J. J. (2023). Proof of concept: Using ChatGPT to teach emergency physicians how to break bad news. *Cureus*, 15(5). <https://doi.org/10.7759/cureus.38755>
- Wee, L. K., Kek, M. Y., & Sim, M. H. (2001). Crafting effective problems for problem-based learning. In *Proceedings of the 3rd Asia-Pacific conference on problem-based learning*.
- Wimmer, M. J., Wilks, D. H., Grammer, R. W., Doerr, R. G., Summers, D. E., & Ressetar, H. G. (2014). Use of patient simulation in problem-based learning for first-year medical students. *Medical Science Educator*, 24, 253–261.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Allison Mooi¹  · Jacob Schmid¹  · Thomas Johnston¹  · Kodi J. Smith McCoy¹  · Zachary Patterson¹  · Hannah Feldt¹  · William Thomas²  · Emma Fenner³  · Kevin Lu¹  · Raya Gandhi¹  · Adam Western¹  · Brendan Seabold¹  · Daniel Vollmer⁴  · Roshan Nallaveetil⁴  · Anthony Fanelli⁴  · Logan Schmillen⁴  · Shelley Tischkau⁵  · Anna T. Cianciolo⁶  · Pinckney Benedict⁷  · Richard Selinfreund⁸ 

✉ Richard Selinfreund
rselinfreund39@siuemed.edu

- ¹ Southern Illinois University School of Medicine, Springfield, IL, USA
- ² Department of Anthropology, Southern Illinois University, Carbondale, IL, USA
- ³ University of Tennessee, Memphis, TN, USA
- ⁴ Department of Computer Science, Southern Illinois University, Carbondale, IL, USA
- ⁵ Department of Pharmacology, Southern Illinois University School of Medicine, Springfield, IL, USA
- ⁶ Department of Medical Education, Southern Illinois University School of Medicine, Springfield, IL, USA
- ⁷ Department of English, Center for Virtual Expression (CVEX), Southern Illinois University, Carbondale, IL, USA
- ⁸ Department of Medical Microbiology, Immunology and Cell Biology, Southern Illinois University School of Medicine, PO Box 19626, Springfield, IL 62794-9626, USA

Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH (“Springer Nature”).

Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users (“Users”), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use (“Terms”). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.

These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.

We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.

While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval, sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.

These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.

To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.

Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.

If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

onlineservice@springernature.com