Powering Clinical Communication Training with an LLM-powered Virtual Reality Simulation

Eileen McGivney
Northeastern University

Xiuqi Tommy Zhu
Northeastern University

Heidi Cheerman
Northeastern University

Sheri Kiami
Northeastern University

Minxin Cheng
Northeastern University

Leanne Chukoskie*
Northeastern University

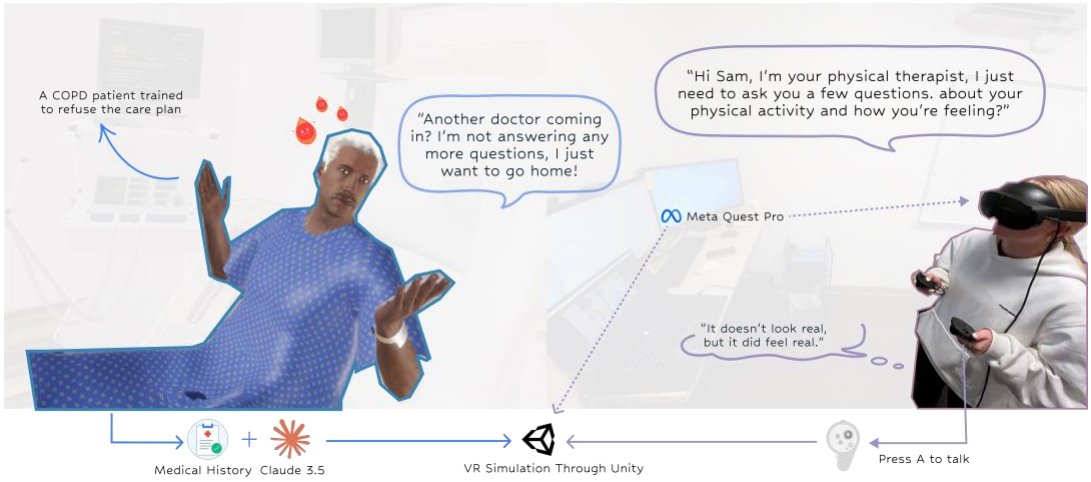

Figure 1: The Virtual AI Patient Simulator (VAPS) enables health profession students to engage in simple, dynamic, voice-based interactions with an LLM-powered embodied virtual patient (left) while immersed in VR using a headset and controllers (right). The figure depicts a challenging patient who resists the proposed care plan, demonstrating how VAPS creates realistic and unpredictable clinical encounters to support authentic learning experiences.

ABSTRACT

Clinical communication is essential for health professions (HP) students, yet simulation labs and prior VR systems often rely on rigid scripts that fail to capture the unpredictability of real patient encounters. We present VAPS, a Virtual AI Patient Simulator that integrates embodied conversational agents (ECAs) with large language models (LLMs) to support dynamic, voice-based interactions. We conducted a mixed-method evaluation with 40 HP students from five disciplines, examining usability, sense of agency, emotional engagement, and perceived realism. Results show that VAPS was usable, reduced nervousness, and was perceived as immersive and emotionally authentic. Students viewed VAPS as a supplement rather than a replacement for simulation labs and envisioned future extensions including diverse patient cases and interprofessional training. Our findings suggest that VR simulation should evolve from looking real toward fostering experiences that feel real, with conversational AI playing a central role in enabling more adaptive and authentic learning.

Index terms: Applied computing → Interactive learning environments.

1 INTRODUCTION

Healthcare professionals, including nurses, physical therapists, physician's assistants, pharmacists, and speech-language pathologists, are increasingly in demand as the United States population grows and ages [1], making health professions (HP) education crucial. For these professions, real-world clinical interaction requires HP students to develop strong interpersonal communication skills in addition to specialized medical knowledge, as well as the ability to interact with patients from diverse backgrounds, needs, and dispositions [2]. Such communication is therefore a life-long essential skill for HP students, but they

* e-mail: l.chukoskie@northeastern.edu

LEAVE 0.5 INCH SPACE AT BOTTOM OF LEFT COLUMN ON FIRST PAGE FOR COPYRIGHT

face challenges acquiring these skills due to the lack of relevant and tailored learning opportunities and poor transfer from training to practice [3]. HP education programs typically prepare students for clinical interactions through classroom activities and simulation-based training that include human standardized patient interactions, allowing for effective hands-on skill learning and confidence building [4].

However, these opportunities are limited by the high cost of equipment, space constraints in universities, scheduling difficulties in traditional clinical simulation, and a lack of authentic hands-on experience through classroom instruction [5]. Recently, Virtual Reality (VR) simulation-based learning has emerged as a promising solution for the creation of engaging, context-specific training environments accessible *whenever and wherever* [6], [7]. Studies show that VR effectively reduces patient safety risk, promotes focused practice, enables students' confidence, and provides contextualized learning experiences [8], [9]. Despite these advantages, however, prior VR clinical simulation systems either overlook patient-provider communication as an important training element [10] or have typically relied on rigid procedures and scripted dialogues for a standardized learning experience [11], [12]. For example, [13] aimed to train clinicians' empathic communication skills by asking participants to select an established response to a patient query in VR scenarios. Such rigidity reduces opportunities for repeated use, limits customizability by educators, and constrains the simulation's capacity to replicate the unpredictability and diversity in real-world patient-provider communication [14], [15].

This research aimed to address the needs of HP students and educators by leveraging unique affordances of VR for training with those of Large Language Models (LLMs) to create dynamic interactions. VR provides opportunities for hands-on training that give learners a great deal of agency to interact with the environment with their full bodies and be in control of their learning [16]. It also situates the user in a first-person perspective, engendering a strong sense of "being there" in the digital environment, engaging learners and improving the transfer of their skills to real-world contexts [7]. Together, these affordances make VR powerful for training, but it remains challenging to design simulations that are authentic to real-world practices and contexts. This study employed a participatory design process to understand the needs of educators and students in HP programs and then developed and tested an LLM-powered virtual reality patient simulator titled VAPS (Virtual AI Patient Simulator), drawing on simulation-based education best practices in HP education [4] and research on embodied conversational agents [17], [18].

2 Formative Study: Participatory Design

This study began with a focus group-based study with faculty from physical therapy, nursing, pharmacy, speech and language pathology, and physician assistant departments that identified interpersonal skills training as a promising area for VR to enhance their students' learning. We then conducted a formative interview-based study to understand student needs and gather design insights that would inform the development of such as system. Student interviews revealed their challenges communicating with patients from diverse backgrounds and varied conditions in their clinical practice, a mismatch between their training opportunities and clinical practice, and positive perceptions of VR for more and more varied practice communicating with patients.

This led to design insights for our system: 1) Design communication training as open and unpredictable conversations; 2) Bring opportunities to allow personalizing simulated scenarios; 3) Reproduce realistic clinical scenarios; 4) Maintain simple interactions that are friendly for VR novices.

3 System Development

From this we designed VAPS as a patient simulator to leverage the dynamic and unpredictable nature of LLMs. Challenging patient cases were drawn from student interviews and clinical faculty expertise, including medical records from a standardized patient case in simulation-based education. The patient, Sam Mason, is hospitalized with an exacerbation of Chromic Obstructive Pulmonary Disease. The patient randomly exhibited one of three challenges: low health literacy, resistant to plan of care, or highly emotional.

The prototype system was developed by the authors using the Meta Quest Pro though it is compatible with any immersive VR headset. The system leverages Unity (Version: 2022.3.61), chosen for its versatility and suitability for immersive VR applications. The virtual patient was created using Character Creator as a high-fidelity model and character detail. We used the Convai API to enable verbal communication, customization, and LLM integration (Claude-3-5-Sonnet). The 3D hospital room scene was constructed using online open-source resources. The character animation was adjusted based on the open-source in Mixamo. Interactions were limited to a push-to-talk button on one controller, and joystick movement on the other. See Figure 1.

4 Exploratory Study: Lab-Based Pilot

In the summer of 2025, we conducted a lab-based pilot of the prototype to explore its potential to simulate meaningful patient interactions and assess student perspectives of practicing clinical communication with an LLM-powered VR simulation. Recruitment was aided by clinical faculty

who identified an appropriate cohort of students who would have training in clinical communication but not yet completed clinical co-ops or rotations. Students were briefed on the objectives of the simulation to conduct a subjective interview or bedside intake. 40 student participated (29 female, 10 male, 1 non-binary; M = 24.35 years; SD = 2.54, range = 21-35). Programs included Physical Therapy (11), Speech and Language Pathology (11), Pharmacy (9), Physician's Assistant (8), and Nursing (1). 35 were graduate students and 5 undergraduates. Participants received a $50 Amazon gift card for participating. This study was approved by the Institutional Review Board (IRB) of Northeastern University.

A pre-survey included demographic questions and the Positive and Negative Affect Schedule (PANAS) [19]. The post-survey also included the PANAS scale, along with measures of sense of presence in VR [20], sense of agency in VR [16], the usability of the VR system (SUS) [21] and three open-ended questions. We conducted a brief semi-structured post-interview focusing on participants' experiences with VAPS.

VAPS received an overall "good" score on the SUS (M = 73.12, SD = 15.19). Students expressed the benefit of minimal interaction demands, which allowed them to concentrate more on crafting their communication approach: "I thought it was intuitive. I mean, I was just kind of talking to him. So I feel like you don't really need to be able to do that much" (P27). On the other hand, students did not feel confident they could use the system on their own: "I don't think I would have been able to do it myself… but once I was in there… it was pretty easy to use" (P11). Interviews also revealed the AI's delayed speech impeded their sense of realism, as the student waited 5-10 seconds for the AI to process what they had said before responding.

Table 1. Results of Pre-Post Affect Scores

	Mean Score						
Measure	**Pre**	**Post**	**Change**	**t**	**p**	**Cohen's *d***	**Sig.**
Positive Affect (Overall)							
Aggregated Scale	**19.05**	**19.60**	0.55	1.94	.060	0.31	ns
Enthusiastic	3.90	3.88	−0.03	−0.23	.822	−0.04	ns
Determined	3.78	4.13	0.35	3.34	.002	0.53	**
Attentive	4.13	4.28	0.15	1.78	.083	0.28	ns
Inspired	3.65	3.68	0.03	0.23	.822	0.04	ns
Excited	3.60	3.65	0.05	0.63	.534	0.10	ns
Negative Affect (Overall)							
Aggregated Scale	**9.70**	**9.10**	−0.60	−2.05	.047	−0.32	*
Nervous	3.03	2.65	−0.38	−3.20	.003	−0.51	**
Afraid	1.93	1.90	−0.03	−0.20	.844	−0.03	ns
Upset	1.38	1.40	0.03	0.19	.850	0.03	ns
Irritable	1.50	1.43	−0.08	−1.14	.262	−0.18	ns
Distressed	1.88	1.73	−0.15	−1.00	.323	−0.16	ns

Note. Significance levels: $*p < .05, **p < .01$.

We observed a small but statistically significant change in students' emotions related to clinical communication from the pre- to post- survey on the PANAS scale, which asked to what degree they felt specific emotions when thinking about communicating with patients (see Table 1). There was a significant decrease in the mean negative emotions (pre: M = 9.70, SD = 0.42; post: M = 9.10, SD = 0.36; p = .047). At the item level, students specifically reported a significant decrease in their level of *nervousness* (t = -3.06, p < .01, Cohen's $d = -0.48). Positive emotions did not significantly change from pre- to post-VAPS (pre: M = 19.05, SD = 0.45; post: M = 19.60, SD = 0.38; p = .06). However, a statistically significant increase in students' feelings of *determination* was observed (t = 3.40, p < .01, Cohen's d = 0.54), indicating an increase in their motivation to communicate well with patients.

Qualitative results also illustrated the emotional engagement students felt, as they described how their interaction engaged emotions they feel in real patient interactions as well. They often described how their emotions changed throughout the interaction: "I think at first, I felt a little nervous when he was kind of abrasive... But then it like I was able to relax and get a little bit more comfortable and develop a rapport with the patient, and I think that I relaxed, and I think he relaxed a little bit more too, because he started to trust me" (P17).

Students also discussed how the unpredictability of the AI simulation contributed to these emotions and helped the VR feel like a real patient interaction: "I do think it's way more realistic than the scripts that our standardized patients have sometimes in class or in laboratory, just because...those concerns of not knowing medical jargon and having to take deep breaths or like pauses and sentences, especially for a disease state, and being really concerned about outside factors...like his dog, wanting his daughter to be there, was more realistic of what a patient would be in clinical setting..." (P25). On the other hand, they expressed how the lack of nonverbal behaviors and quality of the avatar contributed to a sense that the simulation was not real, and expressed how the patient did not look real.

Finally, despite the benefits of an LLM-powered patient to replicate the dynamic and unpredictable nature of patient interactions, student emphasized how VR should only be used as a complement to their existing training, and not as a replacement. They described how having a tool to practice more frequently with varied patients and also to have experience as the sole provider in simulations would be beneficial, in addition to the group exercises they participate in via the physical simulation lab: "I think SIM (simulation) Lab is so inaccessible we only get to do it like once or twice this semester...whereas if you had the opportunity to put on a VR headset once a week, that would be not even once a week, but even every other week it'd be a huge advantage to getting to practice skills more and less of a shock factor when we shift to clinicals" (P17).

5 Conclusion

Combining a formative and pilot study, this project addresses the needs of health profession students to gain more practice opportunities in communication skills than currently available in classroom and simulation lab activities. The formative study helped to understand the needs of HP students and faculty and the most beneficial ways of designing VR to align with their curricula. This surfaced design insights, including the need for simple interactions, dynamic and unpredictable practice opportunities, communicating with challenging patients, and a personalized learning experience. In partnership with clinical faculty who have expertise in simulation-based education, we therefore designed the Virtual AI Patient Simulator, or VAPS, to give students practice opportunities conducting subjective examinations with a challenging patient. The evaluation study assessed the experiences of current students in these programs using VAPS with surveys and interviews, showing that students found the simulation to feel like authentic practice, provide valuable practice opportunities that can augment their current curriculum, and engage important emotions for learning communication skills.

The results highlight how LLMs can generate conversations that feel real, and when implemented in VR the first-person embodied perspective contributes to their sense of being there in a real interaction even with low graphical fidelity. We find that the emotional engagement of students can lead to decreases in negative emotions and may increase motivation to perform well in patient communication.

We are currently further analyzing biometric data and scoring performance of students and the virtual patient to understand how the system can be improved as well as how to automate feedback in the system. Based on the findings from this pilot, we have identified ways to improve the virtual patient including by reducing the AI's speech delay when responding to the student and adding more non-verbal behaviors. In future work we aim to address these issues by building a more robust system with a local AI model. We also plan to create a patient design system that allows clinical faculty to create patient cases for their curricula without needing technical expertise in AI or VR. This will allow VAPS to meet the needs of clinical instructors, who have expressed a desire for patient cases that allow their students to practice on patients with specific conditions, challenges, and diverse identities.

Acknowledgments

The authors wish to thank Michael Bessette, Mollie Moore, Aimee Perron, Alisha Stadnicki, Dawn Swain, Stephen Wood, and Sarah Young-Hmong for their assistance and support of this research. This work was supported by an interdisciplinary seed grant from Northeastern University.

References

[1] Bureau of Health Workforce, "Health Workforce Projections." Accessed: Dec. 19, 2024. [Online]. Available: https://bhw.hrsa.gov/data-research/projecting-health-workforce-supply-demand

[2] R. L. Street, "INTERPERSONAL COMMUNICATION SKILLS IN HEALTH CARE CONTEXTS," in *Handbook of Communication and Social Interaction Skills*, Routledge, 2003.

[3] J. Ammentorp, M. Chiswell, and P. Martin, "Translating knowledge into practice for communication skills training for health care professionals," *Patient Education and Counseling*, vol. 105, no. 11, pp. 3334–3338, Nov. 2022, doi: 10.1016/j.pec.2022.08.004.

[4] S. Kiami and P. Donlan, "The Impact of Simulation Based Experiences on Clinical Education: Physical Therapist Student Perspective," *Journal of Higher Education Theory and Practice*, vol. 22, no. 3, Mar. 2022, Accessed: Dec. 19, 2025. [Online]. Available: https://articlearchives.co/index.php/JHETP/article/view/3363

[5] H. Y. So, P. P. Chen, G. K. C. Wong, and T. T. N. Chan, "Simulation in Medical Education," *Journal of the Royal College of Physicians of Edinburgh*, vol. 49, no. 1, pp. 52–57, Mar. 2019, doi: 10.4997/jrcpe.2019.112.

[6] C. L. Foronda *et al.*, "A Comparison of Virtual Reality to Traditional Simulation in Health Professions Education: A Systematic Review," *Simulation in Healthcare*, vol. 19, no. 1S, p. S90, Jan. 2024, doi: 10.1097/SIH.0000000000000745.

[7] G. Makransky and G. B. Petersen, "The Cognitive Affective Model of Immersive Learning (CAMIL): a Theoretical Research-Based Model of Learning in Immersive Virtual Reality," *Educ Psychol Rev*, Jan. 2021, doi: 10.1007/s10648-020-09586-2.

[8] G. S. Ruthenbeck and K. J. Reynolds, "Virtual reality for medical training: the state-of-the-art," *Journal of Simulation*, vol. 9, no. 1, pp. 16–26, Feb. 2015, doi: 10.1057/jos.2014.14.

[9] C.-H. Chou, H.-C. Tai, and S.-L. Chen, "The effects of introducing virtual reality communication simulation in students' learning in a fundamentals of nursing practicum: A pragmatic randomized control trials," *Nurse Education in Practice*, vol. 74, p. 103837, Jan. 2024, doi: 10.1016/j.nepr.2023.103837.

[10] S. Jeong and C. Cha, "Effects of immersive virtual reality simulation–based maternity nursing education: A randomized controlled trial," *Clinical Simulation In Nursing*, vol. 97, Dec. 2024, doi: 10.1016/j.ecns.2024.101631.

[11] C. Liu, F. F.-Y. Tan, S. Zhao, A. Kanneganti, G. A. Tushar, and E. T. Khoo, "Facilitating Virtual Reality Integration in Medical Education: A Case Study of Acceptability and Learning Impact in Childbirth Delivery Training," in *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, in CHI '24. New York, NY, USA: Association for Computing Machinery, May 2024, pp. 1–14. doi: 10.1145/3613904.3642100.

[12] A. V. Reinschluessel *et al.*, "Versatile Immersive Virtual and Augmented Tangible OR – Using VR, AR and Tangibles to Support Surgical Practice," in *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, in CHI EA '23. New York, NY, USA: Association for Computing Machinery, Apr. 2023, pp. 1–5. doi: 10.1145/3544549.3583895.

[13] D. Sapkaroski, M. Mundy, and M. R. Dimmock, "Immersive virtual reality simulated learning environment versus role-play for empathic clinical communication training," *Journal of Medical Radiation Sciences*, vol. 69, no. 1, pp. 56–65, 2022, doi: 10.1002/jmrs.555.

[14] M. W. Zackoff *et al.*, "Impact of an Immersive Virtual Reality Curriculum on Medical Students' Clinical Assessment of Infants With Respiratory Distress," *Pediatr Crit Care Med*, vol. 21, no. 5, pp. 477–485, May 2020, doi: 10.1097/PCC.0000000000002249.

[15] Q. Sun, J. Huang, H. Zhang, P. Craig, L. Yu, and E. G. Lim, "Design and Development of a Mixed Reality Acupuncture Training System," in *2023 IEEE Conference Virtual Reality and 3D User*

Interfaces (VR), Mar. 2023, pp. 265–275. doi: 10.1109/VR55154.2023.00042.

[16] E. McGivney, "Interactivity and identity impact learners' sense of agency in virtual reality field trips," *Brit J Educational Tech*, vol. 56, no. 1, pp. 410–434, Jan. 2025, doi: 10.1111/bjet.13513.

[17] T. Bickmore and J. Cassell, "Relational agents: a model and implementation of building user trust," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, in CHI '01. New York, NY, USA: Association for Computing Machinery, Mar. 2001, pp. 396–403. doi: 10.1145/365024.365304.

[18] J. Cassell, M. Stone, and H. Yan, "Coordination and context-dependence in the generation of embodied conversation," in *Proceedings of the first international conference on Natural language generation - Volume 14*, in INLG '00. USA: Association for Computational Linguistics, June 2000, pp. 171–178. doi: 10.3115/1118253.1118277.

[19] D. Watson, L. A. Clark, and A. Tellegen, "Development and validation of brief measures of positive and negative affect: The PANAS scales," *Journal of Personality and Social Psychology*, vol. 54, pp. 1063–1070, 1988, doi: 10.1037/0022-3514.54.6.1063.

[20] E. Han *et al.*, "People, places, and time: a large-scale, longitudinal study of transformed avatars and environmental context in group interaction in the metaverse," *Journal of Computer-Mediated Communication*, vol. 28, no. 2, p. zmac031, Mar. 2023, doi: 10.1093/jcmc/zmac031.

[21] john Brooke, "SUS: A 'Quick and Dirty' Usability Scale," in *Usability Evaluation In Industry*, CRC Press, 1996.