Designing VR Simulation System for Clinical Communication Training with LLMs-Based Embodied Conversational Agents

Xiuqi Tommy Zhu zhu.xiu@northeastern.edu Northeastern University United States

Sheri Kiami s.kiami@northeastern.edu Northeastern University Boston, United States Heidi Cheerman h.cheerman@northeastern.edu Northeastern University Boston, United States

Leanne Chukoskie l.chukoskie@northeastern.edu Northeastern University Boston, United States Minxin Cheng cheng.min@northeastern.edu Northeastern University Boston, United States

Eileen McGivney* e.mcgivney@northeastern.edu Northeastern University Boston, United States

in Computing Systems (CHI EA '25), April 26-May 1, 2025, Yokohama, Japan. ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3706599.3719693

1 INTRODUCTION

Healthcare professionals, including nurses, physical therapists, physician's assistants, and pharmacists, are increasingly in demand as the United States population grows and ages [19], making health profession (HP) education crucial. These professions require frequent direct interaction with patients with diverse needs, backgrounds, and dispositions, which require strong interpersonal skills in addition to specialized medical knowledge. To hone these skills, HP education programs typically prepare students through simulation-based training with mannequins or standardized patient interactions with actors, which are effective for teaching handson skills and building confidence through realistic practice [12]. However, simulation-based learning experiences using mannequins and standardized patients face challenges. They can be resourceintensive and costly, and universities have space constraints and challenges in schedule coordination, limiting the opportunities students have to engage in these learning experiences [26].

Virtual Reality (VR) is a promising solution to provide more simulation-based training opportunities in HP education, as its heightened immersion and novel full-body interactivity can create highly engaging training environments tailored to specific needs and used "whenever and wherever" [8, 16, 20]. Studies show that VR training reduces patient risks, enables frequent practice, and provides contextualized learning experiences [23], and that medical VR training simulations are effective across various disciplines, including dentistry [11, 24], surgery [9, 14], and anatomy [25]. Additionally, studies find VR improves students' knowledge acquisition and performance, such as a VR-based childbirth delivery training system that increased learners' knowledge by 24.9% over traditional mannequin-based methods [14].

However, previous VR medical simulation training systems have typically utilized rigid procedures and fixed content in order to provide a standardized learning experience [14, 17, 22]. Such rigidity reduces the potential for repeated use and limits the ability for customization by educators, raising practical concerns for their implementation beyond controlled lab environments [10]. Additionally, previous VR medical simulation training systems often rely on scripted dialogues or step-by-step guidance, limiting their ability to replicate the unpredictability and diversity of real-world medical

ABSTRACT

VR simulation in Health Professions (HP) education demonstrates huge potential, but fixed learning content with little customization limits its application beyond lab environments. To address these limitations in the context of VR for patient communication training, we conducted a user-centered study involving semi-structured interviews with advanced HP students to understand their challenges in clinical communication training and perceptions of VR-based solutions. From this, we derived design insights emphasizing the importance of realistic scenarios, simple interactions, and unpredictable dialogues. Building on these insights, we developed the Virtual AI Patient Simulator (VAPS), a novel VR system powered by Large Language Models (LLMs) and Embodied Conversational Agents (ECAs), supporting dynamic and customizable patient interactions for immersive learning. We also provided an example of how clinical professors could use user-friendly design forms to create personalized scenarios that align with course objectives in VAPS and discuss future implications of integrating AI-driven technologies into VR education.

CCS CONCEPTS

• Human-centered computing → Virtual Reality; • Applied computing → Interactive learning environments.

KEYWORDS

Virtual Reality, Embodied Conversational Agents, Clinical Simulation, Medical Education

ACM Reference Format:

Xiuqi Tommy Zhu, Heidi Cheerman, Minxin Cheng, Sheri Kiami, Leanne Chukoskie, and Eileen McGivney. 2025. Designing VR Simulation System for Clinical Communication Training with LLMs-Based Embodied Conversational Agents. In *Extended Abstracts of the CHI Conference on Human Factors*

CHI EA '25, April 26-May 1, 2025, Yokohama, Japan

© 2025 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-XXXX-X/18/06

https://doi.org/10.1145/3706599.3719693

^{*}Corresponding Author

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

scenarios [24, 25, 27, 31]. Moreover, past VR medical simulation training systems have not considered the needs of students [28], resulting in systems with low student acceptance [14] and a high learning curve [21].

Thus, this work aims to design a VR medical simulation training system that allows dynamic interaction between HP students and simulated patients utilizing a user-centered design method. As part of a larger study on integrating VR into clinical HP programs, this VR system focuses on interpersonal communication skills in clinical settings, a challenge that clinical faculty described their students face due to limited practice opportunities with authentic patient populations. Research also shows interpersonal communication is a life-long essential skill for HP students, but they face challenges due to the lack of relevant and tailored learning opportunities and poor transfer from training to practice [1].

To design a VR simulation system that addresses these challenges, we conducted semi-structured interviews with advanced HP students who had experience working in a clinical setting through internships or residencies to understand their 1) challenges in communication training and clinical practice and 2) perceptions and expectations in VR simulated communication training. Students described a mismatch between their school-based training and clinical practice, particularly due to wide variation in patient conditions, dispositions, and backgrounds. Given these findings, we are currently co-designing VAPS, a Virtual AI Patient Simulator, with clinical professors, which integrates Large Language Models (LLMs) with Embodied Conversational Agents as patients in a hospital setting. Our design leverages the capabilities of LLMs and ECAs to facilitate dynamic interactions and support customizing patients' personalities based on real medical history, meeting learning goals across many HP programs [5]. In this late-breaking work paper, we describe the design of this system to date, along with the insights learned from the user-centered design process. We discuss the implications of integrating LLMs into VR medical simulation training. Given the nature of late-breaking work, we also discuss plans for future iterations of the system and comprehensive evaluation of VAPS

This work makes a number of contributions to the field of VR in Education 1) We provide valuable design insights from Health Professions students for clinical communication training in VR. 2) We present a novel VR clinical simulation system, *VAPS*, with customizing LLMs-based embodied conversational agents and scenarios. 3) We identify implications from the user-centered design process for future VR+LLMs in medical education.

2 INTERVIEW STUDY

As part of a larger project on integrating VR into clinical HP education at a selective US university, we conducted an interview study aiming to answer two research questions: 1) What are HP students' current practices and challenges in clinical communication training? 2) What are their perceptions of using VR for clinical communication training?

2.1 Participants

Our participant criteria required students who had undergone clinical communication training and had at least one instance of on-site clinical communication practice through internships or residencies. Students were recruited by asking HP clinical faculty to refer students and share recruitment materials via email. We conducted interviews with six students (3 Female, 3 Male; M = 25.33 years; SD = 2.21) from three majors: Physician Assistant (PA), Doctor of Pharmacy (PharmD), and Doctor of Physical Therapy (DPT), designated as participants P1 through P6. All participants were familiar with VR in terms of knowing what it is, but there were only two with more extensive use. Participants received a \$10 Amazon gift card for participating. This study was approved by the Institutional Review Board (IRB) of Northeastern University.

2.2 Method and Analysis

Semi-structured interviews consisted of two main parts: (1) experiences in clinical communication training and practice and (2) perspectives on VR for clinical communication training. See Appendix A for the full protocol. Each interview lasted an average of 38 minutes (ranging from 35 to 45). Interviews were conducted in English and were video-audio recorded with participants' consent. Interviews were transcribed and analyzed using the thematic analysis approach [4]. The first author manually reviewed the transcripts for accuracy. Subsequently, two authors independently generated initial codes from all the transcripts using open coding. The authors then discussed their interpretations to reach a consensus and grouped similar codes into clusters. Emerging common themes were identified based on internal connections, and final themes were collaboratively generated using an affinity mapping approach [18].

2.3 Findings

2.3.1 Personalization Challenges during Communication to Diverse Patients. Our participants identified several challenges during the interviews, all converging on the critical issue of personalization in communication. Students described feeling unprepared for some patient interactions due to the patient's condition, such as pain or visible injuries, emotional state, and diverse backgrounds. For example, one student described difficulty in the early days of their internship because they were not used to seeing patients in severe conditions hooked up to complex machines, as one DPT student said: "...I think I learned, I got very acclimated to it, and I very much enjoyed it, but when I walked in, [...] it's my first day, [...] and I saw somebody with end-stage renal disease, and I was there on my first day, that was not something I was prepared for at all...(P5)" Some other students described challenges interacting with patients who were aggressive or refused to engage with them: "...Also, in certain settings, the patient could just be like, really overwhelmed, and have like difficulty answering some questions, or there may be like really sensitive topics, that they don't feel comfortable answering, so it can be hard to get information from them in those cases...(P1)" Our participants also described the diverse backgrounds that patients come from as posing challenges, including treating patients who do not speak English, necessitating an interpreter, or those with low levels of medical literacy, for whom they would need to adapt how they explain conditions and treatments. Overall, students described how patients' conditions, dispositions, and backgrounds varied widely in the contexts they worked in, requiring a great deal of

adaptability. Personalizing their communication for all their patient needs was a challenge for them, particularly at the beginning of their clinical work.

2.3.2 Mismatch between Training and Clinical Communication. When asked about the classroom-based training they received for clinical communication, students pointed to limited opportunities to practice authentic communication, and described a mismatch between their training and what they needed to do in practice. Many students described learning on the job and through their clinical experience as the best way to learn the interpersonal communication skills they needed for their profession. While they did receive training on communication in their coursework, they often identified mismatches between that training and what they needed to do in clinical settings. For example, students described simulation lab activities where they interacted with a standardized patient as beneficial practices, but noted it was insufficient because they only used this lab once or twice a year. Furthermore, they described how their practice opportunities were not always realistic in terms of the interactions they would have with patients. For instance, one DPT student described needing to communicate with people unfamiliar with medical terms versus classroom activities where they practiced with their peers who have advanced medical knowledge, as she stated: "...So [in the lab practical], we're like communicating with our patients. But our patients are classmates. So I feel like just being put in that situation where you have to talk to like an actual patient for that case...(P6)". Students also frequently described patients who did not speak English but did not receive training in communicating through interpreters. As one PA student noted: "...or the patient doesn't really understand what the interpreter saying, or we can't get the exact language that the patient speaks (P1)...".

They also described the ways traditional classroom activities were not as helpful due to the evaluation systems. Students described being focused on getting a good grade and following checklists prescribed by their instructors that added pressure to their simulation and role-play activities in classes and interfered with their ability to practice authentically and flexibly, as one PharmD student said: "... When we do the simulation labs and just like sit in the corner and like ... you know that they're (faculty) looking. You know that they're listening to you, and you're trying to just like drown them out, which can make you feel more nervous because you start to think like I need to make sure I'm saying everything to make sure I'm hitting all the points (P4)..." Grading pressure rendered the training environment less reflective of real clinical settings and, in their view, less meaningful or practical. For example, one PharmD student shared: "...I think just working with real people, you have a little bit more of a curveball with it. And you just like ... everything on that checklist that I used in school, I don't really use that, because it's just different scenarios that are going on...(P3)"

2.3.3 VR as an Opportunity for More Realistic and Varied Practice. During our interviews, most students expressed positive perceptions regarding increased practice opportunities. Notably, participants highlighted the potential of VR as an additional tool for diverse and varied practice scenarios, particularly for extreme situations that are challenging to replicate in standard simulations but are often encountered in real practice. For example, one PA student remarked: "...It could really be used for like so many different things [...] like, what do I do if a patient starts crying in the middle of a conversation, or how do I bring up bad news or things like that?..(P1)". Furthermore, we identified their expectations for various interactive features in VR, including unpredictable conversational dialogues, patient data visualizations, opportunities for self-reflection, and multi-user communication and collaboration. Overall, students identified realistic scenarios and cases to recreate real clinical situations as the most essential factor in VR communication training.

2.4 Design Insights

The themes from the interviews that students face challenges personalizing communication with diverse patients, the mismatch between clinical practice and the classroom training they receive, and the potential of VR to provide more realistic practice opportunities led us to four design insights that meet these goals. In particular, students' discussion of a need for more practice with various patients informs the VR design by prioritizing realistic clinical scenarios with a high degree of personalization and unpredictability over any kind of complex interactivity. We extracted the following design insights: (DI1)Reproduce realistic clinical scenarios: The most important requirement for our VR system is to ensure that the simulation scenarios are as realistic as possible. From the environment to the patients, the system should feature high-fidelity characters, objects, and animations that accurately simulate real-world actions in VR. (DI2)Maintain simple interaction to create easy learning experiences: To help students gain essential skills and confidence, our system should prioritize meaningful VR content over complex or unrealistic interactive features. Since HP students rarely have opportunities to engage with VR, reducing their learning effort through an intuitive, easy-to-use system will enhance their overall training experience. (DI3)Bring opportunities to allow personalizing simulated scenarios: To meet more diverse learning goals, our VR system should have the ability to support customization for a wide range of patient personas and clinical contexts. Thus, users can reuse an identical system that is customized and changed through different learning objectives each time. (DI4)Designing communication training as open and unpredictable conversations: When simulating communication training in VR, it is essential to go beyond step-by-step, scripted learning approaches. Open and unpredictable conversations reflect the human nature of real interactions, as patients rarely follow a fixed script, therefore making users feel they are engaging with real patients.

3 VAPS: VIRTUAL AI PATIENT SIMULATOR

3.1 System Overview

Driven by the above design insights, we design a novel Virtual reality system empowered by conversational **AI** Patient Simulator: *VAPS*. The design of this system draws on both the insights identified from the student interview study as well as best practices in HP education simulation, which both emphasize using VR as a practice opportunity to apply prior knowledge, not a didactic instructional system. The system therefore has three scenes to simulate the entire process: tutorial, clinical patient interaction, and reflection. The goal of this system is to enhance HP students' clinical communication

skills by providing realistic and dynamic practice opportunities. In the *tutorial* scene, users are introduced to interacting with the VR system. They are guided through interactive guidelines to explore and familiarize themselves with the interaction of *VAPS* with the help of a high-fidelity ECA "tour guide." In this scene, the learner is also pre-briefed on the patient case and learning objectives. In the *clinical patient interaction* scene, we designed another ECA to act as a patient lying in a simulated hospital room (Fig.1). Learners enter the room and must sit near the patient's bed to collect their history and details about their current condition. After this interaction, learners must identify a referral or consultation for the patient, then enter the final *reflection* scene. In this scene, students reflect on their performance based on standardized prompts and align with simulation education standards.

3.2 Embodied Conversational Agents

Inspired by our DI1. DI3 and DI4, we determined the capabilities of Large Language Models (LLMs) would be the optimal solution to address the need for dynamic and customizable conversations, and chose ECAs as the vehicle to deploy LLMs. This aligns with previous research findings, which suggest that combining non-verbal conversational and empathetic cues-such as hand gestures and facial expressions-with spoken dialogue can significantly enhance user trust in ECAs [3]. Thus, our ECAs have the following features informed by design insights, 1) Personalized ECAs Design: Our ECAs can be trained from prompts in the patient's medical history and personal backstory. Also, users can choose their personality traits, state of mind, preferred language, and speaking style during the design process. 2) Integration with Medical Records: Our ECAs support uploading real medical files to a knowledge repository, ensuring accurate and context-sensitive responses. 3) Dynamic Narrative Design: Our ECAs can be designed with narrative flow. Users could design ECAs to adapt to any triggers or decisions during interactions. 4) Realistic Animations: Our ECAs perform basic head movement and hand gestures, as well as synchronized lip motions, to create human-like, realistic interactions during communication.

The design of the clinical patient interaction scenario draws on best practices in simulation for HP education, and we are codesigning this with a clinical professor with over 15 years of experience designing clinical simulation training. While the prototype begins with one scenario and a limited number of patient personas for learners to practice (see Section 3.4), ultimately, the system will allow clinical faculty to customize the simulation based on their course objectives and learner needs. To address the challenges faced by non-AI experts, such as clinical professors, in creating prompts for ECAs' every feature, we developed a user-friendly design form (See Appendix B). For example, clinical professors can input patient backstories that are common materials for simulation exercises and select levels for the patient's personality in the design form, which is then converted into a patient within the VAPS system by the development team. This also emphasizes the importance of reviewing scripts manually to ensure the appropriateness of training prompts. Finally, the LLMs trained for VAPS incorporate patient profile information into profile-specific embodied character models.

3.3 System Interaction

Our interaction design, inspired by *DI2*, emphasizes simplicity for ease of learning. Thus, users can engage with the system in *VAPS* through three primary actions: Move Joystick (to teleport to a specific location), Push Button (to talk, inspect, or reset), and Press Trigger (to interact with buttons in the VR environment). Each button has a distinct purpose, enabling versatile interactions. For example, the right controller's 'A' button enables voice communication, while the 'B' button shows a patient information board with key details like medical history and demographics. These straightforward interactions are designed to minimize learning barriers and enhance usability.

3.4 Example Learning Case

The initial prototype utilizes one learning case that illustrates how VAPS can be integrated into HP education. In this scenario, the learner must conduct a subjective interview with a patient who is diagnosed with chronic obstructive pulmonary disease (COPD), a progressive lung disease. This diagnosis is a good candidate for the prototype because professionals from all the target clinical programs would need to treat a patient with this disease, namely nurses, physician's assistants, physical therapists, and pharmacists. Our ECAs allow learners from these varied professions to conduct the relevant patient interview, which may be focused more on physical activity or medications. The patient's diagnosis and condition remain constant for the simulation, but the patient persona will vary according to four common challenges students face: 1) a highly emotional patient, 2) an aggressive or rude patient, 3) a patient who is refusing care, and 4) a patient with low medical literacy. Learners will therefore be able to practice conducting the subjective interview multiple times to adapt to different patient conditions. The use of LLMs allows the interaction to vary dynamically in each use, and the reflective activity helps them identify areas of improvement between each round of practice.

The proposed implementation would allow students to practice with *VAPS* on their own, for example, homework, and then have a reflective activity with their instructor or a small group of peers. This allows for more in-depth debriefing, aligning with standard practices in simulation, as typically conducted following an in-person simulation session. This learning case serves as just an illustrative example rather than a prescriptive guide for implementing *VAPS* in actual learning contexts, but ultimately, *VAPS* will facilitate repeated use by putting design in the hands of instructors. Clinical professors can modify the scenario and the patient personas to introduce different learning objectives aligned with their courses through the design form. We discuss more learning scenarios and the vision of *VAPS* in Section 4.2.

3.5 Implementation

The prototype system was developed by the research team using Meta Quest Pro and Meta Quest 3, though it is compatible with any immersive VR headset. The system leverages Unreal Engine 5, chosen for its versatility and suitability for immersive VR applications. The ECAs were created using MetaHuman¹ as a high-fidelity

¹https://www.unrealengine.com/en-US/metahuman

Designing VR Simulation System for Clinical Communication Training with LLMs-Based Embodied Conversational Agents CHI EA '25, April 26-May 1, 2025, Yokohama, Japan



a) Tutorial Scene





c) Debriefing Scene

Figure 1: Overview of VAPS: a)Freely-explore tutorial scene; b)A High-fidelity realistic ECA is communicating with users as an AI patient in clinical patient interaction scene; c) Quick self-debriefing science with simple open-questions

model and character detail. We used Convai Api² to enable communication, customization, animation, and LLMs integration (GPT-40, generally regarded as the state-of-the-art LLM) employed in VAPS. The 3D hospital room scene was constructed using online opensource resources. For the reflection and tutorial rooms, we utilized the standard Unreal VR template scenarios, prioritizing the content within these steps over the environmental details.

DISCUSSION 4

Supporting Customized VR Learning with 4.1 LLMs

Our interview findings highlight challenges health professionals face when communicating with highly dynamic patients in clinical settings, which students reported they need more practice on before entering clinical practice. This informs a key design implication for VR simulation systems: the need to support customizable scenarios that address diverse learning objectives. Previous literature has highlighted the potential of LLMs in educational and social VR environments [15, 29] as well as the advancement of ECAs in clinical VR systems [6, 13]. Building on these insights, we combine these two elements-LLMs and ECAs-into the design of a VR simulation to support communication skills practice in a more dynamic clinical setting. Our co-design process also led to the identification of LLMs as a powerful tool to address a felt need in health professions education of customization need, rather than simply following the current "LLMs + Everything" trend. VAPS employs LLM-based ECAs as direct objects of human-AI interaction to facilitate learning rather than mediating interactions between humans or non-intelligent agents. Compared to VR, other mobile applications such as computer screens or iPads, may offer greater accessibility in clinical training and require less technical expertise to interact with LLMs. However, given the use of ECAs as vehicles and the focus on communication skills as a learning goal in this paper, we utilize VR because these aspects can be simulated more naturally in an immersive environment. Future research could compare the specific performance of different vehicles in this context.

An important feature of VAPS is its flexibility. While the case of communicating with a patient diagnosed with COPD described in Section 3.4 will be our initial test, the system allows for instructors to determine other patient conditions and upload their medical history. The personas of the patients are also customizable to provide

students with varied practice with realistic patients. This flexibility makes VAPS stand out from previous VR medical simulation systems, as it allows students to encounter the unpredictability and diversity of real-world medical scenarios while enabling faculty to customize different learning goals without rigid instructional constraints. Further, the implementation of the tool is adaptive to instructors' needs. VAPS could be used before or after in-person simulation experiences to provide additional practice opportunities. It may be used as homework by individual students who then debrief their performance in class, or instructors may ask small groups of students to use it in class. In the future, we aim to add more features to VAPS that will broaden its applications, including outputs like video and text summaries that can enable reflection and instructor feedback, multiplayer modes where students can practice interprofessional communication that include other healthcare professionals, and options for communicating via interpreters when treating patients who are not fluent in English. With these capabilities, VAPS will enhance customization meeting various needs in clinical communication training.

We also recognize concerns about AI ethics and data privacy. LLMs have a tendency to hallucinate and reproduce biases from their training data. Also, user data when interacting with an LLM is not always secure. VAPS addresses these issues by putting the training of the AI patient in the control of the research team along with clinical faculty. Through rigorous training we can specify what LLMs can and cannot include in its interactions and test it for bias. In our future user study, we will assess any risks participants face from bias and hallucination. In the future, our goal is to use a closed system that will keep data private rather than a commercial product.

4.2 Leveraging VR to Expand Realistic and **Dynamic Training Opportunities**

Our interview findings align with Yarmand et al. [30], indicating a mismatch between training and clinical practice. Merely replicating traditional simulation practices in VR risks reproducing this mismatch. However, our findings illustrate how students perceive VR as a promising medium for providing realistic and dynamic practice opportunities. Therefore, VAPS focuses on dynamic training procedures that mimic unpredictable situations encountered in real clinical practice, enhanced by high-fidelity animation, textures, and models. Due to most students and faculty being unfamiliar with

²https://convai.com/

VR, *VAPS* prioritizes simple interactions and environments to avoid distracting learners or making them inaccessible.

During the development of our system and the training of ECAs, we encountered both the challenges and importance of persona development, as suggested by Chin et al. [7]. VAPS addresses these challenges by helping clinical professors design prompts, something that non-AI experts typically struggle with [32]. Doing so required collaboration with clinical faculty to design the ECAs and the development of a form that simplifies the LLMs training process. Our system allows faculty to contribute domain-specific information and features that increase realism including narrative flow, medical records, and personality metrics. This promotes dynamic practice that is aligned with the curriculum and addresses limitations in traditional VR clinical simulations, which often rely on a rigid and fixed learning structure [14, 17, 22]. However, we acknowledge our system has some limitations. These include a focus primarily on communication training and a fixed VR environment, which restricts the ability to simulate diverse clinical scenarios, such as a physical therapist treating patients in rural hospitals. Thus, we advocate that future work should not only build upon this foundation to support more clinical scenarios but also support interprofessional education, such as enabling multiple HP students from various disciplines to study together. Additionally, with advancements in generative AI for 3D modeling [33] and human animation [2], we are optimistic that these constraints can be addressed in the future. However, bringing VR to higher educational institutions ultimately requires efforts from the entire network of stakeholders [10].

5 CONCLUSION AND FUTURE WORK

In this paper, we explored the design of LLM-based ECAs for VR clinical communication training. We began with an interview study involving six Health Profession students to examine current clinical communication training practices and their perceptions of VR training. Based on the findings, we proposed a set of design insights and developed a prototype VR system, *VAPS*, to address the identified design requirements. Furthermore, we highlighted the significance of customization and realism in VR simulation systems. For future work, we plan to evaluate *VAPS* in the real learning context. Our evaluation roadmap includes conducting an initial usability test followed by long-term studies to assess the effectiveness and performance of VR-based simulation training in comparison to traditional simulation training, specifically in the domain of clinical communication.

ACKNOWLEDGMENTS

We are grateful for all the students and faculty who participated in this research. We are also grateful to Ara Jung from the ReGameXR Lab for her assistance. This research was supported by a Northeastern University TIER 1 interdisciplinary research grant.

REFERENCES

- Jette Ammentorp, Meg Chiswell, and Peter Martin. 2022. Translating knowledge into practice for communication skills training for health care professionals. *Patient Education and Counseling* 105, 11 (Nov. 2022), 3334–3338. https://doi.org/ 10.1016/j.pec.2022.08.004
- [2] Samaneh Azadi, Akbar Shah, Thomas Hayes, Devi Parikh, and Sonal Gupta. 2023. Make-An-Animation: Large-Scale Text-conditional 3D Human Motion

Generation. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). 15039–15048.

- [3] Timothy Bickmore and Justine Cassell. 2001. Relational agents: a model and implementation of building user trust. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Seattle Washington USA, 396–403. https://doi.org/10.1145/365024.365304
- [4] Virginia Braun and Victoria Clarke. 2012. Thematic analysis. American Psychological Association.
- [5] Justine Cassell. 2000. Embodied Conversational Agents. MIT Press. Google-Books-ID: tHiKZGh9t7sC.
- [6] Vuthea Chheang, Shayla Sharmin, Rommy Márquez-Hernández, Megha Patel, Danush Rajasekaran, Gavin Caulfield, Behdokht Kiafar, Jicheng Li, Pinar Kullu, and Roghayeh Leila Barmaki. 2024. Towards Anatomy Education with Generative AI-based Virtual Assistants in Immersive Virtual Reality Environments. In 2024 IEEE International Conference on Artificial Intelligence and eXtended and Virtual Reality (AIxVR). 21–30. https://doi.org/10.1109/AIxVR59861.2024.00011 ISSN: 2771-7453.
- [7] Jessie Chin, Smit Desai, Sheny (Cheng-Hsuan) Lin, and Shannon Mejia. 2024. Like My Aunt Dorothy: Effects of Conversational Styles on Perceptions, Acceptance and Metaphorical Descriptions of Voice Assistants during Later Adulthood. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW1 (April 2024), 1–21. https://doi.org/10.1145/3637365
- [8] Cynthia L. Foronda, Laura Gonzalez, Merrick M. Meese, Nicholas Slamon, Mariju Baluyot, Jiye Lee, and Michelle Aebersold. 2023. A Comparison of Virtual Reality to Traditional Simulation in Health Professions Education: A Systematic Review. Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare (Aug. 2023). https://doi.org/10.1097/SIH.000000000000745
- [9] Danilo Gasques, Janet G. Johnson, Tommy Sharkey, Yuanyuan Feng, Ru Wang, Zhuoqun Robin Xu, Enrique Zavala, Yifei Zhang, Wanze Xie, Xinming Zhang, Konrad Davis, Michael Yip, and Nadir Weibel. 2021. ARTEMIS: A Collaborative Mixed-Reality System for Immersive Surgical Telementoring. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. ACM, Yokohama Japan, 1–14. https://doi.org/10.1145/3411764.3445576
- [10] Qiao Jin, Yu Liu, Svetlana Yarosh, Bo Han, and Feng Qian. 2022. How Will VR Enter University Classrooms? Multi-stakeholders Investigation of VR in Higher Education. In CHI Conference on Human Factors in Computing Systems. ACM, New Orleans LA USA, 1–17. https://doi.org/10.1145/3491102.3517542
- [11] Maximilian Kaluschke, Rene Weller, Myat Šu Yin, Benedikt W. Hosp, Farin Kulapichitr, Siriwan Suebnukarn, Peter Haddawy, and Gabriel Zachmann. 2024. Reflecting on Excellence: VR Simulation for Learning Indirect Vision in Complex Bi-Manual Tasks. In 2024 IEEE Conference Virtual Reality and 3D User Interfaces (VR). IEEE, Orlando, FL, USA, 712–721. https://doi.org/10.1109/VR58804.2024.00091
- [12] Sheri Kiami and Pamela Donlan. 2022. The Impact of Simulation Based Experiences on Clinical Education: Physical Therapist Student Perspective. *Journal of Higher Education Theory and Practice* 22, 3 (March 2022). https://articlearchives. co/index.php/JHETP/article/view/3363 Number: 3.
- [13] Liliana Laranjo, Adam G Dunn, Huong Ly Tong, Ahmet Baki Kocaballi, Jessica Chen, Rabia Bashir, Didi Surian, Blanca Gallego, Farah Magrabi, Annie YS Lau, et al. 2018. Conversational agents in healthcare: a systematic review. *Journal of the American Medical Informatics Association* 25, 9 (2018), 1248–1258.
- [14] Chang Liu, Felicia Fang-Yi Tan, Shengdong Zhao, Abhiram Kanneganti, Gosavi Arundhati Tushar, and Eng Tat Khoo. 2024. Facilitating Virtual Reality Integration in Medical Education: A Case Study of Acceptability and Learning Impact in Childbirth Delivery Training. In Proceedings of the CHI Conference on Human Factors in Computing Systems. ACM, Honolulu HI USA, 1–14. https://doi.org/10.1145/3613904.3642100
- [15] Ziyi Liu, Zhengzhe Zhu, Lijun Zhu, Enze Jiang, Xiyun Hu, Kylie A Peppler, and Karthik Ramani. 2024. ClassMeta: Designing Interactive Virtual Classmate to Promote VR Classroom Participation. In Proceedings of the CHI Conference on Human Factors in Computing Systems. ACM, Honolulu HI USA, 1–17. https: //doi.org/10.1145/3613904.3642947
- [16] Guido Makransky and Gustav B. Petersen. 2021. The Cognitive Affective Model of Immersive Learning (CAMIL): a Theoretical Research-Based Model of Learning in Immersive Virtual Reality. *Educational Psychology Review* 33, 3 (Sept. 2021), 937–958. https://doi.org/10.1007/s10648-020-09586-2
- [17] Thomas Muender, Anke Verena Reinschluessel, Daniela Salzmann, Thomas Lück, Andrea Schenk, Dirk Weyhe, Tanja Döring, and Rainer Malaka. 2022. Evaluating Soft Organ-Shaped Tangibles for Medical Virtual Reality. In CHI Conference on Human Factors in Computing Systems Extended Abstracts. ACM, New Orleans LA USA, 1–8. https://doi.org/10.1145/3491101.3519715
- [18] Michael Muller. 2014. Curiosity, Creativity, and Surprise as Analytic Tools: Grounded Theory Method. In Ways of Knowing in HCl, Judith S. Olson and Wendy A. Kellogg (Eds.). Springer New York, New York, NY, 25–48. https: //doi.org/10.1007/978-1-4939-0378-8_2
- [19] Bureau of Health Workforce. 2024. State of the U.S. Health Care Workforce, 2024. https://bhw.hrsa.gov/sites/default/files/bureau-health-workforce/state-ofthe-health-workforce-report-2024.pdf

Designing VR Simulation System for Clinical Communication Training with LLMs-Based Embodied Conversational Agents CHI EA '25, April 26-May 1, 2025, Yokohama, Japan

- [20] Jack Pottle. 2019. Virtual reality and the transformation of medical education. Future Healthcare Journal 6, 3 (Oct. 2019), 181–185. https://doi.org/10.7861/fhj. 2019-0036
- [21] Manuel Rebol, Krzysztof Pietroszek, Neal Sikka, Claudia Ranniger, Colton Hood, Adam Rutenberg, Puja Sasankan, and Christian Gütl. 2023. Evaluating Augmented Reality Communication: How Can We Teach Procedural Skill in AR?. In 29th ACM Symposium on Virtual Reality Software and Technology. ACM, Christchurch New Zealand, 1–11. https://doi.org/10.1145/3611659.3615685
- [22] Anke Verena Reinschluessel, Thomas Muender, Roland Fischer, Valentin Kraft, Verena Nicole Uslar, Dirk Weyhe, Andrea Schenk, Gabriel Zachmann, Tanja Döring, and Rainer Malaka. 2023. 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. ACM, Hamburg Germany, 1–5. https://doi.org/10.1145/3544549.3583895
- [23] G S Ruthenbeck and K J Reynolds. 2015. Virtual reality for medical training: the state-of-the-art. *Journal of Simulation* 9, 1 (Feb. 2015), 16– 26. https://doi.org/10.1057/jos.2014.14 Publisher: Taylor & Francis _eprint: https://doi.org/10.1057/jos.2014.14.
- [24] Sara Samuel, Carmine Elvezio, Salaar Khan, Laureen Zubiaurre Bitzer, Letty Moss-Salentijn, and Steven Feiner. 2024. Visuo-Haptic VR and AR Guidance for Dental Nerve Block Education. *IEEE Transactions on Visualization and Computer Graphics* 30, 5 (May 2024), 2839–2848. https://doi.org/10.1109/TVCG.2024.3372125
- [25] Danny Schott, Matthias Kunz, Tom Wunderling, Florian Heinrich, Rüdiger Braun-Dullaeus, and Christian Hansen. 2023. CardioGenesis4D: Interactive Morphological Transitions of Embryonic Heart Development in a Virtual Learning Environment. *IEEE Transactions on Visualization and Computer Graphics* 29, 5 (May 2023), 2615–2625. https://doi.org/10.1109/TVCG.2023.3247110 Conference Name: IEEE Transactions on Visualization and Computer Graphics.
- [26] Hing Yu So, Phoon Ping Chen, George Kwok Chu Wong, and Tony Tung Ning Chan. 2019. Simulation in Medical Education. *Journal of the Royal College of Physicians of Edinburgh* 49, 1 (March 2019), 52–57. https://doi.org/10.4997/jrcpe. 2019.112
- [27] Qilei Sun, Jiayou Huang, Haodong Zhang, Paul Craig, Lingyun Yu, and Eng Gee Lim. 2023. Design and Development of a Mixed Reality Acupuncture Training System. In 2023 IEEE Conference Virtual Reality and 3D User Interfaces (VR). IEEE, Shanghai, China, 265–275. https://doi.org/10.1109/VR55154.2023.00042
- [28] Wai Tong, Meng Xia, and Huamin Qu. 2024. Exploring Stage Lighting Education in Metaverse. In Extended Abstracts of the CHI Conference on Human Factors in Computing Systems. ACM, Honolulu HI USA, 1–6. https://doi.org/10.1145/ 3613905.3650924
- [29] Hongyu Wan, Jinda Zhang, Abdulaziz Arif Suria, Bingsheng Yao, Dakuo Wang, Yvonne Coady, and Mirjana Prpa. 2024. Building LLM-based AI Agents in Social Virtual Reality. In Extended Abstracts of the CHI Conference on Human Factors in Computing Systems. 1–7.
- [30] Matin Yarmand, Chen Chen, Kexin Cheng, James Murphy, and Nadir Weibel. 2024. T'd be watching him contour till 10 o'clock at night': Understanding Tensions between Teaching Methods and Learning Needs in Healthcare Appenticeship. In Proceedings of the CHI Conference on Human Factors in Computing Systems. ACM, Honolulu HI USA, 1–19. https://doi.org/10.1145/3613904.3642453
- [31] Matthew W Zackoff, Francis J Real, Rashmi D Sahay, Lin Fei, Amy Guiot, Corinne Lehmann, Ken Tegtmeyer, and Melissa Klein. 2020. Impact of an Immersive Virtual Reality Curriculum on Medical Students' Clinical Assessment of Infants With Respiratory Distress. *Pediatric Critical Care Medicine* 21, 5 (2020).
- [32] J.D. Zamfirescu-Pereira, Richmond Y. Wong, Bjoern Hartmann, and Qian Yang. 2023. Why Johnny Can't Prompt: How Non-AI Experts Try (and Fail) to Design LLM Prompts. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (Hamburg, Germany) (CHI '23). Association for Computing Machinery, New York, NY, USA, Article 437, 21 pages. https://doi.org/10.1145/ 3544548.3581388
- [33] Lei Zhang, Jin Pan, Jacob Gettig, Steve Oney, and Anhong Guo. 2024. VRCopilot: Authoring 3D Layouts with Generative AI Models in VR. In Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology (Pittsburgh, PA, USA) (UIST '24). Association for Computing Machinery, New York, NY, USA, Article 96, 13 pages. https://doi.org/10.1145/3654777.3676451

A INTERVIEW PROTOCOL

Part One: Interviewee Background (5-10 minutes).

- Can you tell me what program you are in? What year are you in that program?
- Why did you choose to enroll in this program in college?
- What health professions courses have you taken at (Blinded) University?

- What are your experiences and knowledge about Virtual Reality (VR)?
- What clinical training and practice did you receive?
- Part Two Interview Questions (30-40 minutes).
 - Tell me about any training you have received in communication in clinical settings, like delivering information to patients or having conversations with them. What did you learn? How was the training conducted?

a. During the training, what challenges did you face with the current practices, and why?

b. Is there anything in your training process that you think can be improved?

• In your work and studies, do you need to collaborate with other clinicians from different disciplines? Have you received any training about that?

a. If yes, what challenges have you encountered when you collaborate with other professions?

b. How do you balance communicating with patients and other clinicians?

c. If no, can you imagine what happens when you would need to work with a clinician from another discipline?

• What do you think is the most helpful method for learning how to communicate in clinical settings?

a. During your communication with patients, which aspects would you value more?

• What are your thoughts on how virtual reality could be integrated into your education for training on communication in clinical settings?

a. If positive, what expectations do you think virtual reality can help achieve that traditional methods cannot in your learning? And what interactive features you would like to use?

b. If not, why do you think this is not feasible and what kind of effort is needed?

• Is there anything else that you would like to contribute that we have not discussed?

B DESIGN FORMS

Fill out this form to design an AI patient in VSPA! Medical History:

- Reason for Visit:
- Past Medical History:
- Family Medical History:
- Current Medications:
- Allergies:
- Smoking:
- Alcohol Use:
- Exercise Habits:
- Diet:

Information:

- Name:
- Age:
- Gender Identity:
- Ethnicity/Race:
- Language(s) Spoken:

CHI EA '25, April 26-May 1, 2025, Yokohama, Japan

- Accent:
- Occupation:

Personal backstory:

Describe the character's backstory as instructions in second person i.e. phrases like "You are.., You are an expert on.., Your life is based on the following facts..." Try to keep the backstory limited to a short background of the character and the world / scenario they are present in, basic personality and mannerisms. For longer factual information, use the knowledge bank.

Personality

score with 1-5 for each characteristic, see Fig 2 Narrative Design (Optional)

Event l

Event Objective Character Action Trigger

• Event 2

Event Objective Character Action Trigger

- Event 3
 - Event Objective Character Action Trigger



Figure 2: A user-friendly design form for helping non-AI experts design the features and characters of AI patients in *VAPS*

C EXAMPLE PROMPTS OF TRAINING ECAS BY CLINICAL PROFESSOR

Reason for Visit: Admitted to the hospital 2 days ago following a severe COPD exacerbation triggered by a respiratory infection. Symptoms included severe shortness of breath, wheezing, productive cough, and low oxygen saturation (SpO2 85% in air in the room).

Past Medical History:

- Chronic obstructive pulmonary disease (COPD) diagnosed 7 years ago.
- Hypertension controlled with medication.
- Type 2 diabetes mellitus managed with oral agents.

- Anxiety mild, occasional symptoms managed without medication.
- History of recurrent bronchitis.

Family Medical History:

- Father: Passed away at 68 from lung cancer (smoker).
- Mother: History of type 2 diabetes and hypertension.
- Brother: Diagnosed with COPD, current smoker.

Current Medications:

- Albuterol inhaler (as needed).
- Fluticasone/salmeterol (Advair Diskus) twice daily.
- Tiotropium (Spiriva) once daily.
- Metformin 1,000 mg twice daily.
- Lisinopril 10 mg daily.

Allergies:

• Sulfa drugs (causes rash and swelling).

Smoking:

• 50 pack-years (smoked 1 pack per day for 50 years). Quit 2 years ago after his first hospitalization for COPD exacerbation.

Alcohol Use:

Occasionally drinks beer on weekends, 1-2 bottles per occasion.

Exercise Habits:

 Minimal activity due to breathlessness but enjoys short walks around his neighborhood when able.

Diet:

• High in carbohydrates (e.g., rice, beans, tortillas). Loves fried foods and sugary drinks. Attempts to follow dietary recommendations for diabetes but struggles with portion control.

Information:

- Name: Samuel "Sam" Rivera
- Age: 62
- Gender Identity: Male (he/him/his)
- Ethnicity/Race: Hispanic/Latino
- Language(s) Spoken: Primarily Spanish; conversational English.
- Accent: Noticeable Spanish accent with a deliberate and slow speech pattern in English.
- Occupation: Retired auto mechanic.

Personal backstory:

Sam is a retired auto mechanic who worked in the same shop for over 35 years. He takes great pride in his work and often reminisces about the cars he fixed and the friends he made. He lives in a small apartment with his dog, Luna, who is his primary companion. His wife passed away 10 years ago from breast cancer, and he has two adult children who live out of state. Sam identifies as a devout Catholic and often prays for strength when facing challenges. He struggles with feelings of loneliness, particularly since his children visit infrequently. Despite this, he maintains a positive outlook most of the time and is well-liked by his neighbors, who occasionally check in on him. He speaks Spanish at home and prefers to communicate in Spanish, although he can manage in English for basic conversations. He expresses frustration when medical instructions are not clear or when he feels rushed during interactions. Sam is Designing VR Simulation System for Clinical Communication Training with LLMs-Based Embodied Conversational Agents CHI EA '25, April 26-May 1, 2025, Yokohama, Japan

cooperative with medical staff but admits to feeling overwhelmed by his condition and the numerous medications he needs to take. He often relies on his daughter, who lives in Florida, for emotional support over the phone.

Personality

- Openness 2
- Meticulous 3
- Extraversion 2
- Agreeableness 3
- Sensitive 3

Narrative Design

Event l Event Objective

- Demonstrate empathy and active listening.
- Provide appropriate emotional support.
- Collaborate with other team members to address mental health needs.
- Character Action: Crying and Upset Patient

Trigger:

During session, Sam becomes tearful and expresses feelings of hopelessness about his condition, saying, "I can't do this anymore. I'm a burden to everyone." Key Points for Each Discipline:

- Physical Therapy: Pause the session, provide reassurance, and adjust the plan to match his emotional and physical capacity.
- Nursing: Explore underlying emotional concerns, offer resources for mental health support, and coordinate with the care team.
- Pharmacy: Discuss the potential need for anxiolytic or antidepressant therapy.
- SLP: Assess if emotional distress impacts swallowing safety and ability to participate in therapy.
- PA: Validate his concerns, provide support, and potentially consult psychiatry or counseling services.

Event 2 Event Objective

- De-escalate the situation calmly and professionally.
- Maintain boundaries while addressing the patient's concerns.
- Recognize potential triggers and underlying issues for anger.
- Character Action Angry and Inappropriate Behavior

Trigger:

Sam becomes agitated during a nursing assessment, yelling, "You people don't know what you're doing! This place is useless!" He makes a derogatory comment toward a staff member. Key Points for Each Discipline:

- Nursing: Use a calm tone, acknowledge his frustration, and attempt to identify specific complaints.
- Physical Therapy: Adjust the session as needed, ensuring safety and addressing barriers to cooperation.
- Pharmacy: Assess if medications could contribute to irritability (e.g., corticosteroids).
- SLP: Evaluate whether difficulty communicating or swallowing might be contributing to his frustration.
- PA: Lead a debrief with the care team to align approaches and discuss possible psychiatric or behavioral interventions.

Event 3 Event Objective

- Communicate effectively with a patient who is reluctant or overwhelmed.
- Assess and adapt health education for low medical literacy.
- Encourage engagement in care while respecting autonomy.
- Character Action Refusing Care and Low Medical Literacy

Trigger:

Sam refuses to attend therapy sessions and states, "What's the point? My lungs are shot anyway." He also struggles to understand instructions for his inhaler, stating, "I don't need all these fancy gadgets." Key Points for Each Discipline:

- Physical Therapy: Use motivational interviewing techniques to explore barriers and encourage participation.
- Nursing: Simplify health education materials and provide hands-on demonstrations for self-management.
- Pharmacy: Explain medication benefits in simple, relatable terms, and demonstrate inhaler techniques.
- SLP: Ensure that instructions are clear and accessible, considering potential swallowing or cognitive barriers.
- PA: Address his refusal empathetically, emphasizing achievable goals and the importance of adherence to the care plan.

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009