City Research Online

City St George's, University of London

Citation: Reimers, S. & Myers, L. (2024). Using generative AI agents for scalable roleplay activities in the health sciences. In: Beckingham, S., Lawrence, J., Powell, S. & Hartley, P. (Eds.), Using Generative AI Effectively in Higher Education: Sustainable and Ethical Practices for Learning, Teaching and Assessment. (pp. 82-86). London: Routledge. ISBN 9781032773988 doi: 10.4324/9781003482918-12

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/33418/

Link to published version: https://doi.org/10.4324/9781003482918-12

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to City Research Online policy.

City Research Online: http://openaccess.city.ac.uk/ publications@citystgeorges.ac.uk

DRAFT CHAPTER FOR INCLUSION IN

Beckingham, S. (Ed.) *Using Generative AI effectively in HE: sustainable and ethical AI for the common good*. Routledge

Title: Using Generative AI agents for scalable roleplay activities in the health sciences

Keywords: Roleplay, interprofessional learning, group activities, experiential learning, active learning, simulation.

Introduction

Artificial intelligence is being incorporated rapidly into medical and health sciences practice and training (for a review see Abd-Alrazaq et al., 2023). In professional practice it can be used to make more accurate and efficient diagnoses and support medical decision making (see He et al., 2019); in training it can be used in simulation to develop students' clinical and communication skills (see Stamer, Steinhäuser & Flägel, 2023).

We examined the use of Generative AI (GenAI) in roleplay activities where an AI Agent – here, an AI simulation of a character – takes the role of a key participant in a scenario. Roleplay has a significant part in skills training for health professionals and can be a useful experiential learning tool to develop students' interpersonal and professional skills. Although professional actors are sometimes used, more frequently all roles – patient and clinician – are played by students, which can feel irrelevant or embarrassing, particularly for students playing patients (Stevenson & Sander, 2002). It also requires careful and effortful facilitation. The potential benefits of roleplay are, however, significant. Students get to experience a range of realistic situations and develop communication and judgement skills, flexibility, and confidence, from the interactions.

We looked at using GenAI to play the role of a patient in a complex case scenario. The rationale was twofold: First, to avoid the awkwardness and challenges of students having to play the role of a patient; Second, to allow scalability so that multiple groups could engage in the roleplay with limited facilitation.

The aim was not to have a real-time conversation with the AI patient, but to create space to unpack a professional interaction, pausing to explore and discuss potential courses of action at each stage of the conversation with input from students of different disciplines, and subsequently seeing the outcome of those strategies.

Case Study: Interprofessional Learning

Interprofessional learning in the health sciences, where several professions learn from and about each to improve professional working, is increasingly included in undergraduate training. We recently developed a series of events where undergraduate students of midwifery, nursing, and speech and language therapy participated in small, interdisciplinary groups, working through activities together.

In one day-long scenario, students engaged with a fictional case study involving a woman, Sarah, who was 38 weeks pregnant at the start of the scenario, with a nine-year-old son who had communication difficulties. The scenario developed over the course of the day, with a number of traditional discursive and analytical activities; the AI-facilitated roleplay

activity reported here occurred where Sarah had been discharged after giving birth and was at home being visited by the Speech and Language therapist.

We built a simple GPT-based roleplay web app using the React and node.js JavaScript frameworks, where groups of students interacted with an AI agent that played the role of Sarah. Building this required a working knowledge of JavaScript and the relevant frameworks, although there are now AI-supported non-coding alternatives, such as OpenAI's freestanding GPTs (https://openai.com/blog/introducing-gpts). The app took typed natural language user input entered into a browser; combined it with a hidden 250-word system prompt providing background and the current scenario, along with any preceding conversation, and high-level instructions for how to respond; and passed it to the OpenAI GPT3.5 API, displaying the response when received.

The 45-minute session comprised 15 groups of 6-10 students, with one facilitator working across several groups. Students were given scenario details, and then asked to introduce themselves to 'Sarah' as if they were knocking on her front door. They saw Sarah's response on the screen – in a similar format to a WhatsApp chat – and discussed what approach to take in the next message. The session proceeded in a self-paced manner. In most cases, agreed responses were typed by one student on a computer connected to a projector, although some groups used their own mobile phones.

The system prompt, which gave instructions to the GPT regarding the context, Sarah, and how to respond to different approaches, was not seen by students, so although they knew the details of the scenario, and could see their own submissions and responses from Sarah, they did not know everything about Sarah's current state of mind or preferences. The system prompt was designed to generate discussion about different courses of action: For example, Sarah would not allow the speech and language therapist into her home, so as the interaction developed students discussed whether to be more persistent, try a different approach, or refer to a different agency.

As Sarah's support needs emerged during the roleplay, students discussed how to handle them, and through which professional agency, learning about the resources and support each profession was able to provide as they progressed.

Experience

Although the timescale for implementing the activity meant that we were unable to formally examine student perceptions formally, we were present in around six of the sessions ourselves, where we spoke to around 20 students informally during and after the activity; we also spoke to facilitators of other sessions about their experience.

Feedback based on these informal discussions and observations was generally positive. Unlike traditional roleplay, students showed little embarrassment; the whole group was involved and needed little supervision. Students could choose how involved they wanted to be in discussing and typing; we did not encounter any students who did not want to participate at all. Students liked the open-ended nature of the interaction: They could try different things – ask to hold the baby, ask about the son's wellbeing – and the AI agent almost always gave appropriate responses, and several described an emotional engagement with 'Sarah' (one midwifery student saying "I just want to give her a hug"). Students liked working together as a group, having to respond flexibly to developments, and learning about each other's professions.

There was some variability across groups. Where there was less discussion, the task could be completed fairly quickly: One facilitator reported that their conversation fizzled out because of the AI Sarah was overly compliant, agreeing with the majority of suggestions. The scenario was also fairly static and lacking in explicit goals so felt a little aimless to some. A more structured, explicitly goal-directed activity could have been more effective overall. Although students liked the novelty of the interaction with an AI agent, they sometimes focused on testing the limits of the system rather than trying to make the best possible responses.

Future development and broader applications

Learning from this initial experience, adjustments to both the prompt and the web app could substantially improve the activity for future students. One development would be to use some simple logic within the web app to change the system prompt fed to the GPT agent: For example, the patient should become distressed or aggressive after a certain period; a new character could arrive at a particular juncture. We are also looking at using the GPT agent to provide automated constructive feedback at the end of the interaction, related to the professionalism of the conversation, the options explored, and overall outcome. We plan to repeat this activity using an updated version of the web app and gathering more formal feedback from students.

We have focused on health sciences, but the potential for AI-based roleplay in higher education is much broader. We have recently started working with our Business School to use a similar approach to negotiation skills training, giving students the opportunity to explore different negotiation strategies with an AI agent before or after applying them to classroom activities. One advantage of using a simple web app to interact with the GPT API is that scenarios can be tweaked or completely changed by editing the system prompt, without the need to alter the code. There is the potential for easily using GenAI for training for a range of professional interactions and decisions with clients, service users, colleagues, and others.

Wider Caveats

In developing our activities we have encountered potential issues with AI-based roleplay that align with more general concerns about use of GenAI. The first is around the type of character that the AI presents: By default, the system communicates articulately and eloquently, able to explain its own motivation and reasoning, which would usually not capture the range of people that the training would usually seek to build skills with. Essentially by default it could end up training people to interact effectively with a particular type of introspective, well-educated native English speaker. It is easy enough to add instructions to a prompt specifying the socioeconomic status, neurodiversity, or other properties of the character for the AI to play but then there is the risk that it presents a caricature. This is something that merits scenario-level examination before use.

There are also potential risks around accuracy and bias when the AI agent integrates wider information into its response. In our scenario when asked questions beyond its briefing – the son’s communication skills and wellbeing, or specifics about the home environment – the system usually gave sensible responses that fitted the scenario, as an actor might do. However, when drawing on knowledge outside that given by the scenario there is again the risk that in addition to potentially including inaccurate information the response reflects its training data, and hence predominantly experiences of white, affluent westerners rather

than the range of real patients' experiences (for a broader discussion see Karabacak et al., 2023).

More generally we would not recommend replacing existing roleplay or simulated patient activities with AI-based interactions. The experience of interacting with a real person – fellow student, actor, or patient – has many advantages in terms of realism and engagement. However the scalability, flexibility, and repeatability of AI-based activities mean they can complement face-to-face training by allowing students to explore a wider variety of scenarios, alone or with peers, and consider different approaches to addressing them.

Critical Appraisal

Charanjit Sangha

BSc Speech and Language Therapy Student; City, University of London

I was a participant in the chatbot patient activity. I thought it was quite novel and modern. It was engaging, and I liked how you couldn't completely predict what it was going to do next or how it would respond to your actions and suggestions, which made it more like the reality of those situations. It made you think on your feet – a useful professional skill – and got us all involved in the task. It encouraged teamwork in deciding roles and pulling in different professions when it looked like they were needed.

We weren't sure of the end goal, and sometimes ended up in cul-de-sacs where the chatbot paused and we weren't sure what to do next. This was quickly overcome when we realised that we could restart it and try again. It was the more tech savvy people who led initially, but everyone did join in.

I thought it might work better with smaller groups, maybe groups of four rather than our group of eight, with one person from each profession so everyone has to contribute. Having smaller groups completing the activity and then evaluating their different approaches could have been interesting.

Overall I thought it was an enjoyable, useful activity. It generated an upbeat, informal discussion at the end about what different groups had tried and where they ended up. I think this kind of activity can definitely be run again – particularly with more detailed instructions and aims.

References

Abd-Alrazaq, A. et al. (2023) Large Language Models in Medical Education: Opportunities, Challenges, and Future Directions. *JMIR Medical Education, 9(1),* p.e48291. Available at: https://doi.org/10.2196/48291.

He, J., et al. (2019) The practical implementation of artificial intelligence technologies in medicine. *Nature Medicine, 25(1),* pp.30-36.

Karabacak, M. et al. (2023). The Advent of Generative Language Models in Medical Education*. JMIR Medical Education, 9,* p.e48163. Available at: https://doi.org/10.2196/48163.

Stamer, T., Steinhäuser, J. and Flägel, K., 2023. Artificial Intelligence Supporting the Training of Communication Skills in the Education of Health Care Professions: Scoping Review. *Journal of Medical Internet Research, 25*, p.e43311. Available at: https://doi.org/10.2196/43311.

Stevenson, K. and Sander, P., 2002. Medical students are from Mars-business and psychology students are from Venus-University teachers are from Pluto? *Medical Teacher, 24(1),* pp.27-31.

Stian Reimers, Professor of Psychology and Behavioural Science and Co-Director for Educational Enhancement and Digital Innovation, School of Health and Psychological Sciences, City, University of London. stian.reimers@city.ac.uk.

Lucy Myers, Senior Lecturer in Speech and Language Therapy and Co-Director for Educational Enhancement and Digital Innovation, School of Health and Psychological Sciences, City, University of London. lucy.myers@city.ac.uk.

Stian Reimers is a cognitive psychologist specializing in research on human decision making. He also conducts research and develops technology around pedagogy and the student learning experience.

Lucy Myers is a Speech and Language Therapist who specialises in working with school aged children with language impairment. Her research covers language and literacy difficulties and the life experiences of young adults with language and communication needs.

Abstract

This chapter discusses the use of Generative AI (GenAI) in group roleplay activities for health sciences students' interprofessional learning. A simple GPT-based roleplay web app was developed, allowing multidisciplinary groups of students to interact with an AI agent playing the role of a patient in a complex case scenario. This provided challenges and dilemmas, prompting students to discuss multi-agency handling of support needs. The approach received positive feedback, with high engagement, minimal embarrassment, and emotional connection with the GenAI patient. Limitations such as GenAI compliance and scenario dynamics were identified and discussed, along with potential solutions and broader issues.