12

Training GPT as a Standardized Patient

Mercedes Lorena Pedrajas López
Universidad Antonio de Nebrija, Spain
mpedrajas@nebrija.es
https://orcid.org/0000-0003-4257-9260

Ana Sanz Cortés
Universidad Antonio de Nebrija, Spain
asanzco@nebrija.es
https://orcid.org/0000-0002-5153-5860

Eva García Carpintero-Blas
Universidad Antonio de Nebrija, Spain
egarcibl@nebrija.es
https://orcid.org/0000-0002-4984-2511

Esther Martínez Miguel
Universidad Antonio de Nebrija, Spain
emartinezmi@nebrija.es
https://orcid.org/0000-0002-5153-5860

Sara Uceda Gutiérrez
Universidad Antonio de Nebrija, Spain
suceda@nebrija.es
https://orcid.org/0000-0002-1467-5186

Abstract
The integration of artificial intelligence (AI) into learning environments poses a challenge in advancing towards more efficient interactive methodologies. The use of AI-based learning assistants, especially generative language models like OpenAI's GPT, can expand the scope of methodologies such as clinical simulation by generating interactions where AI assumes the role of a standardized patient. Clinical simulation recreates, substitutes, and/or extends real experi-

ences through guided experiences that evoke or replicate substantial aspects of the real professional context in a fully interactive manner. The standardized patient is an actor trained to perform predefined responses based on the students' behaviour and performance. With appropriate AI training, focused on instruction and adaptation to different patient profiles based on their health-disease processes, it is possible to design and implement clinical simulation scenarios where students interact with it. The authenticity of AI allows achieving a high degree of fidelity, and its scope surpasses the limits of synchronous in-person demand of a standardized actor, exponentially multiplying the capacity to generate simulated learning environments. This chapter outlines the keys to integrating AI as a standardized patient into simulated learning experiences.

Keywords: Artificial Intelligence (AI), standardized patient, clinical simulation.

12.1. The importance of simulation in university education

Simulation in university education is crucial for preparing students in various professional fields. It provides a realistic environment to apply theories and skills acquired in the classroom, promoting critical thinking and teamwork (Gormley et al., 2023). These learning experiences transform the student into the protagonist, with the professor as a guide (Cheng et al., 2016). Simulation evolves with technology, integrating virtual environments and augmented reality to offer immersive experiences that prepare students for future challenges (Kononowicz et al., 2019). Clinical simulation with standardized patients is particularly valuable in the training of healthcare professionals. These patients reproduce physical symptoms and emotions, allowing students to improve diagnostic, communication, and decision-making skills in a safe and educational environment (Abshire et al., 2020). The use of ChatGPT (GPT) as a standardized patient in clinical simulation is relevant in healthcare, offering realistic interactions that enhance communication skills and competencies in patient care, especially in terms of clinical assessment and reasoning. Its versatility allows adapting to various clinical scenarios, enriching the simulation experience. The combination of advanced technology and traditional simulation methods prepares healthcare professionals to face complex scenarios in their future practice. Therefore, the con-

vergence of simulation and technological innovation is key to achieving educational excellence in healthcare.

12.2. ChatGPT as a Tool for Learning

In the current digital era, AI has emerged as a revolutionary tool in education, offering innovative opportunities to enhance the teaching-learning process. AI recreates human intelligence processes through algorithms and systems that analyse massive data to identify patterns, make decisions, and predict outcomes. Unlike traditional programs, it can learn and adapt, enabling machines to perform tasks that previously required human intervention (Rouhiainen, 2018). GPT stands out as a versatile and valuable tool for teachers and researchers seeking to optimize the educational experience (Leng, 2024; Abdellatif et al., 2022; Kung et al., 2022). OpenAI's GPT-4, a natural language-based virtual assistant, has found application in various sectors, including customer service, computing, health, and education. Its ability to process text and adapt to different domains has made it a versatile and valuable tool for optimizing experiences in these fields (Morcela, 2023).

Regarding the use of GPT in teaching health-related subjects, several authors have highlighted that its use has provided students with information quickly, without waiting for responses from teachers, thanks to its ability to gather information on a wide variety of topics to assist students in real-time and available 24 hours a day (Leng, 2024; Abdellatif et al., 2022; Kung et al., 2022).

12.3. Understanding how GPT Works

Once the pillars of clinical simulation and the characteristics of AI are understood, they are proposed to train GPT-4 as a standardized patient.

Before working with GPT-4, it is important to understand how it functions. It is a tool that operates based on language analysis and can comprehend and generate text similarly to how a human would. The essence of its operation lies in machine learning and the ability to anticipate patterns; thus, it has been

trained to analyse and understand the structure and meaning of sentences, adapting to a wide range of queries and commands. Its ability to generate contextually appropriate responses is based on the vast amount of information it has absorbed during its training, so the more it is used, the better it will adapt its responses to the user.

Being a language analysis-based tool, the choice of verbs and the formulation of prompts are fundamental aspects to shape the identity and style of the Chat; therefore, it is advisable to consider the following:

- Verbs: Their choice can influence the tone and style of the generated responses. More formal or informal, active or passive verbs contribute to the construction of GPT's analysis and response pattern. If the characteristics of the character's personality are properly defined, it is recommended to work with verbs that can provide coherence. For example, using expressions with passive verbs like "tell me about..." may imply a "service" pattern for which the Chat is designed, which could break the simulated role of identity. It is suggested to use questions with active verbs such as "what do you think about...", "how do you feel that...", or "give me an example...", and observe which of them generates responses consistent with the desired outcome.
- Formulation of Prompts: Linked to the choice of the verb, it is important that, before using the patient created in a simulation, the prompt that will guide the experience is clearly formulated. The "prompt" is the precise and specific instruction that helps generate the activity, that is, the exact statement of what is expected of the Chat during that period. Once the identity has been worked on, when bringing it into the simulation space, it is important to clearly communicate what is expected of it. For example, it can be said: "Now we will simulate that you are a patient attending a psychological consultation. I will ask you questions as your psychologist to understand what is happening to you, and you must respond as the patient, using the information provided earlier in a creative way." In this case, the proposed role, the working context, the objective of the activity, and how the previously provided information is expected to be used are observed.

- Context Markers: These are keywords that will grant coherence to the character. In the context of a clinical simulation, where a complete life history will be provided, the use of dates, proper names, or locations will help improve identity construction. The responses generated by the Chat will be more consistent with the objectives. Therefore, it is important to consider these markers when designing the simulation briefing, so that students can formulate questions using these keywords.

12.4. Establishing the GPT Identity as a Standardized Patient

It may seem like a significant challenge to give it an identity, considering the tool's functionality, but it will be the time of practice and some specific guidelines that will help with this:

- Define the traits that characterize its personality: this involves having a clear image of the person in question, focusing on the details that will influence its communication style. It is essential to define aspects such as communication tone, topics of interest, level of linguistic formality, age, characteristic expressions or slang, and gender of the character being represented.
- To ensure the consistency of the defined character, it is crucial to develop a chat dedicated exclusively to building that profile. This involves avoiding contradictory instructions to GPT regarding its interactions, questions, or writing style. Suddenly switching to another profile can disrupt its learning process. Therefore, it is recommended to create a new designated chat with the name of the character and dedicate time to interact with him or her to improve consistency.
- It is essential to define the specific context in which the interaction will take place. By understanding this context, potential discrepancies between questions asked and the Chat's responses can be minimized. Therefore, it is fundamental to determine the clinical context that is desired to be simulated this time.
- Finally, in the process of constructing the character's identity, it is crucial to provide feedback to the tool. If during the conversations to develop this identity, the interactions do not meet expectations, it is important to communicate this using

the evaluation function provided along with the response or reconsider the formulation of the question. The model's adaptability is based on patterns and data correlations, so these evaluations are of great value in improving its performance.

Once the identity is defined, it is necessary to design the clinical context that frames the standardized patient being worked with. To ensure the credibility of this profile, it is ideal to provide GPT with detailed case information.

In this regard, since the interaction occurs through the exchange of electronic messages, it may not be necessary to consider the phenotypic characteristics of the patient, but demographic characteristics should be considered. These variables, along with the medical history, symptoms, and signs relevant to the case and of clinical interest, should be well defined. All of this facilitates the AI having all the necessary information to contextualize the case and achieve the most natural and coherent behavior possible throughout the entire interview or interaction.

Some other issues to consider in defining the patient profile refer to how the AI should behave during the interaction with the student. It is important to provide clear instructions on the purpose of the interaction and the role that GPT will play in it; these instructions should be consistent with standardized patient training in clinical simulation methodologies (Ruiz & Caballero, 2014).

Therefore, it is recommended to:

- Instruct GPT to speak in detail about the relevant topics for the case without providing any clues necessary for the student's approach or resolution.
- Request that it behaves naturally throughout the interaction.
- Emphasize the need for GPT to internalize its role as a standardized patient and 'feel' as if it were actually the person/patient being portrayed.
- Indicate to GPT the need for consistency in the message throughout the interaction that takes place in the exchange of electronic messages.

On the other hand, the design of standardized responses for a simulated patient is crucial for the simulation success, as it di-

rectly influences the student's performance, similar to interacting with a real patient during clinical practice. These standardized responses will include information on emotional communication, clinical evolution, and shared patient decisions processes, enabling the student to effectively apply their skills to achieve the health goals set for the patient. It is essential to provide a detailed and coherent description of the health problem and its context, which will stimulate the student´s interpersonal and clinical skills development and facilitate their training.

It is relevant to note that AI used as a simulated patient does not possess consciousness or emotions, and although it simulates a conversation, it lacks the deep understanding and empathy characteristic of human intelligence. Its operation is based on patterns and statistical correlations to provide relevant responses, which means that, at times, its responses may not be as expected. Therefore, to define the role that AI assumes as a standardized patient, it is necessary to establish the corresponding context and narrative. Additionally, the profile of the student with whom the AI will interact should be considered, taking into account their level of knowledge, maturity, and expected learning outcomes.

The design of logical interaction is relatively simple, but the difficulty lies in training responses that fit the context, situation, and personality of the standardized patient when the student deviates from what is expected. In this sense, creating scripted behavioural trees representing the sequence of responses for each case will be helpful.

12.5. How to Integrate GPT as a Standardized Patient in the Curricular Design

Given the potential impact of GPT, academics must integrate it in the whole teaching-learning process, including curricula design, learning outcomes, methods and assessment methodologies (McCoy et al., 2020).

Learning outcomes

To maximize the effectiveness of simulated situations and comprehensively address student development, it is important to

consider both the cognitive and emotional aspects of the proposed practice when presenting the objectives in agreement:

- Cognitive objectives should consider both the knowledge and skills for the correct development and resolution of the situation presented through AI.
- Affective objectives encompass the abilities to establish an empathetic relationship even though the standardized patient is AI, those aimed at assessing the difficulties of the simulation situation and enhancing learning capacity, and those related to reflection and self-efficacy.

Below is a guide table for formulating these learning outcomes in clinical simulation (Table 12.1):

Table 12.1. Learning outcomes in simulation scenarios with AI as a standardized patient

	Type of objetive	Description
Cognitive	Knowledge in clinical practice	Applies theoretical concepts in a practical context. Demonstrates a deep understanding of the simulated pathology or condition.
	Technical skills	Integrates technical procedures with acquired theoretical knowledge. Demonstrates specific technical skills related to the simulated clinical situation.
	Decision making	Makes clinical decisions based on available information. Evaluates and prioritizes healthcare-clinical issues effectively.
	Communication	Demonstrates appropriate communication skills with the standardized patient. Adapts communication to the environment in which the experience takes place.
Affective	Empathy and sensitivity	Shows empathy towards the simulated patient, acknowledging their feelings and concerns. Demonstrates sensitivity to cultural and social diversity.
	Stress and pressure management	Manages stress and pressure in the clinical situation. Handles the difficulty of using the chat as part of the simulated situation (developing resilience).
	Reflection and self-evaluation	Reflects on own practice and seeks opportunities for improvement. Provides constructive evaluations to peers.

Formative rubric

Based on the learning outcomes, the design of a formative rubric to assess student performance in simulation with AI as a standardized patient is necessary to provide measurable and specific feedback (Adamson et al., 2013). For its proper development, the following points need to be considered:

- Define learning outcomes: This aspect involves not only identifying cognitive and affective objectives, as previously indicated, but also ensuring they are aligned with the overall simulation goal.
- Identify evaluation criteria in activities that are measurable and specific. Depending on the scope of application as a standardized patient, these may include clinical knowledge and reasoning, or effective communication. Empathy and sensitivity should always be present in the exercise.
- Establish performance levels, each of which should have clear and specific descriptions. An example would be: unsatisfactory, basic, competent, and outstanding.
- Develop levels` description. It should be as detailed as possible and objectively recorded through the written information in the chat, which the teacher will later have access to.
- Structure the rubric according to categories that establish the relative weight of performance, according to the importance of each.
- Include a space for specific teacher comments that encompass both strengths and areas for improvement.
- Seek feedback from other teachers or clinical-healthcare professionals to ensure the validity of the rubric. Adjustments may be made if needed.

A formative rubric example for the affective outcome of maintaining empathetic communication with the standardized patient would be:

- Insufficient (1): The student demonstrates poor communication with the patient, showing no empathy or recognition of their concerns or feelings.
- Basic (2): The student communicates adequately with the patient, although improvement in empathy and clarity could be made.

- Competent (3): The student demonstrates effective communication with the patient, showing empathy and using clear and understandable language.
- Excellent (4): The student establishes exceptional communication with the patient, demonstrating empathy, understanding, and effectively adapting to the patient's needs.

A formative rubric example for the cognitive objective of identifying symptomatology in the first session of clinical assessment is provided below:

- Insufficient (1): The student fails to correctly identify the symptoms presented by the patient.
- Basic (2): The participant identifies some symptoms of the patient but omits important details or makes errors in the description.
- Competent (3): The student identifies the majority of the patient's symptoms accurately and comprehensively, providing detailed and relevant descriptions.
- Excellent (4): The student accurately identifies all of the patient's symptoms, providing thorough descriptions and demonstrating an understanding of the clinical importance of each.

Simulation scenario preparation

Before the scenario begins, students should receive a briefing or prebriefing that provides them with sufficient context to aim for specific learning objectives. This introduction is defined by the Agency for Healthcare Research and Quality as the guidance that provides students with preparatory instructions and information before the simulation session (Lioce et al., 2020).

It may include the creation of a safe container (Rudolph et al., 2014; Turner & Harder, 2018), which is a learning space free from judgment where all participants commit to reality and fiction (commitment from instructors to recreate reality and commitment from students to abstract from the rest and focus on fiction as a recreation of that reality), confidentiality (information about what happened is not shared with students from other courses or groups), and competence (all participants are intelligent, competent, and strive to learn).

With or without a safe container, the briefing involves providing students with key information to address the simulation scenario, and in the case of simulation with AI as a standardized patient, it should include some specific considerations. Since it is a digital or telemedicine consultation, it is important to indicate to students that they must always maintain the professional role they assume in the simulation, taking care of their interaction with the same considerations as with a real patient in face-to-face consultation or intervention. The interpersonal distance in digital health care may be increased by simulation with AI as a standardized patient, due to the interface and the type of language that the AI uses. In this sense, it will be necessary to emphasize the importance of not losing the reality-fiction commitment with the AI. To facilitate AI responses tailored to the proposed clinical case for which it has been trained, students must also consider the information provided in the fourth key, regarding verbs and prompts.

Debriefing

Debriefing is the phase where knowledge is constructed. Students, guided by the instructor, reflect on the action and build knowledge, generating a framework of shared thinking where error is an ally and reflective practice (Schön, 1992) is promoted. Furthermore, discussion of thinking frameworks and the search for and proposal of alternative actions or solutions in future situations are encouraged (Díaz & Cimadevilla, 2019). Reviewing the simulated episode by analyzing actions and reflecting on the role of thought processes, psychomotor skills, and emotional states in them allows maintaining and improving performance in the future. In general, and regardless of the model used, debriefing can be structured into three phases (Motola et al., 2013):

- Reaction Phase: This phase is aimed at dissipating cognitive obstacles and generating a conducive framework for analysis by exploring and verbalizing students' emotions during the simulation, seeking to prevent them from interfering with cognitive processes. In addition, an atmosphere conducive to reflective conversation should be generated, and individual and collective discourse should be harmonized around what happened during the simulation.

- Analysis Phase: This phase focuses on reflection, on analysis for the search and construction of mental models that explain why the student acted the way they did, through guided inquiry questions from the instructor, which, far from providing standard solutions, stimulate doubt and controversy, thus guiding reflection.
- Summary Phase: The experience is synthesized, generating a shared mental framework of possible actions and decisions in future actions, and the session is closed with some dynamics that allow distilling key concepts of the knowledge built in the session.

Evaluating the experience

Self-efficacy is a framework that can help determine students' learning experience when this methodology is applied. This is defined as the individual's perception of their ability to successfully perform a task (Bandura, 1977). Although this variable is not a precise indicator of the level of learning, it is defined as a predictor of the student's success in achieving it (Barrios et al., 2017). In health sciences, simulation with a standardized patient is considered more effective than traditional training for improving learning-related self-efficacy (Merchán-Baeza et al., 2021).

Therefore, it is recommended to use a measurement instrument with optimal psychometric qualities validated for the target population. An example could be the General Self-efficacy Scale, created by Schwarzer & Jerusalem (1993).

12.6. Conclusion

The growing influence of artificial intelligence in various areas, including education, highlights the need not only to ensure that students acquire AI literacy but also, given the potential impact of this technology, update curricula to reflect this new landscape. Academics must reflect on how these advancements will affect curriculum design, assessment methodologies, and pedagogical approaches. Therefore, it is crucial to continuously research and evaluate to achieve the effective integration of AI-

based learning tools in education. This challenge raises the following question: How can we productively incorporate the use of AI in classrooms?

In the health education area, clinical simulation with standardized patients emerges as conducive to this integration scenario, which can also benefit from the potential of AI to multiply the scope of clinical simulation learning experiences.

References

Abdellatif, H., Al Mushaiqri, M., Albalushi, H., Al-Zaabi, A. A., Roychoudhury, S., & Das, S. (2022). Teaching, learning and assessing anatomy with Artificial Intelligence: The Road to a better future. *International Journal of Environmental Research and Public Health*, *19*(21), 14209. https://doi.org/10.3390/ijerph192114209

Abshire, M. A., Li, X., Basyal, P. S., Teply, M. L., Singh, A. L., Hayes, M. M., & Turnbull, A. E. (2020). Actor feedback and rigorous monitoring: Essential quality assurance tools for testing behavioral interventions with simulation. *PloS one*, 15(5), e0233538. https://doi.org/10.1371/journal.pone.0233538

Adamson, K., Kardong-Edgren, S. y Willhaus J. (2013). An updated review of published simulation evaluation instruments. *Clinical Simulation in Nursing*, *9*, e393-e400. http://dx.doi.org/10.1016/j.ecns.2012.09.004

Bandura A. (1977). Self-efficacy: Toward a Unifying theory of behavioral change. *Psychological Rev*, *84*(2), 91-215.

Barrios, S., Urrutia, M. y Rubio, M. (2017). Impacto de la simulación en el desarrollo de la autoeficacia y del locus de control en estudiantes de enfermería. *Educación Médica Superior*, *31*(1), 125-136. https://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-21412017000100012&lng=es&tlng=es

Cheng, A., Morse, K. J., Rudolph, J., Arab, A. A., Runnacles, J., & Eppich, W. (2016). Learner-centered debriefing for health care simulation education: Lessons for faculty development. Simulation in healthcare. *Journal of the Society for Simulation in Healthcare*, *11*(1), 32-40. https://doi.org/10.1097/SIH.0000000000000136

Díaz, D. A., & Cimadevilla, B. (2019). Educación basada en simulación: *debriefing*, sus fundamentos, bondades y dificultades. *Simulación Clínica*, *2* (1), 95-103. https://dx.doi.org/10.35366/RSC192F

Gormley, G. J., Carr, D., Murphy, P., Tallentire, V. R., & Smith, S. E. (2023). Unlocking the learning potential of simulation-based education. *British Journal of Hospital Medicine, 84*(12), 1-8. https://doi.org/10.12968/hmed.2023.0353

Kononowicz, A. A., Woodham, L. A., Edelbring, S., Stathakarou, N., Davies, D., Saxena, N., Tudor Car, L., Carlstedt-Duke, J., Car, J., & Zary, N. (2019). Virtual patient simulations in health professions education: Systematic review and meta-analysis by the digital health education collaboration. *Journal of Medical Internet Research, 21*(7), e14676.

Kung TH, Cheatham M, Medenilla A, Sillos C, De Leon L, Elepaño C, et al. (2022). Performance of GPT on USMLE: potential for AI-assisted medical education using large language models. *PLoS Digit Health, 2*(2), e0000198.

Leng, L. (2024). Challenge, integration, and change: GPT and future anatomical education. *Medical Education Online, 29*(1), 2304973. https://doi.org/10.1080/10872981.2024.2304973

Lioce L. (Ed.), Lopreiato J. (Founding Ed.), Downing D., Chang T.P., Robertson J.M., Anderson M., Diaz D.A., Spain A.E. (Eds.), &Terminology and Concepts Working Group (2020). *Healthcare Simulation Dictionary* (2nd edition). Agency for Healthcare Research and Quality. https://doi.org/10.23970/simulationv2

McCoy, L. G., Nagaraj, S., Morgado, F., Harish, V., Das, S., & Celi, L. A. (2020). What do medical students actually need to know about artificial intelligence? *NPJ Digital Medicine, 3*, 86. https://doi.org/10.1038/s41746-020-0294-7

Merchán-Baeza, J. A., González-Sánchez, M., & Pérez-Cruzado, D. (2021). Simulación con pacientes estandarizados en ciencias de la salud: una revisión sistemática. *Revista Chilena de Terapia Ocupacional, 22*(2), 25-43. https://doi.org/10.5354/0719-5346.2021.61071

Morcela, Ó. A. (2023). GPT: la IA está aquí y nos desafía. *Revista Internacional de Ingeniería Industrial, 6*, 3-6. http://www3.fi.mdp.edu.ar/otec/revista/index.php/AACINI-RIII/article/view/67

Motola, I., Devine, L. A., Chung, H. S., Sullivan, J. E., & Issenberg, S.B. (2013). Simulation in healthcare education: A best evidence practical guide. *Med. Teach., 82*(10) 1511-1530. https://doi.org/10.3109/0142159X.2013.818632

Rouhiainen, L. (2018). *Inteligencia artificial. 101 cosas que debes saber hoy sobre nuestro futuro*. Alienta.

Rudolph, J. W., Raemer, D. B., & Simon, R. (2014). Establishing a safe container for learning in simulation the role of the presimulation

briefing. *Simul Healthc, 9*(6), 339- 349. https://doi.org/10.1097/SIH.0000000000000047

Ruiz, R., & Caballero, F. (2014). Programa para seleccionar y entrenar pacientes estandarizados en el contexto de un currículo universitario de simulación clínica. *Fundación Educación Médica, 17*(4), 199-204. https://dx.doi.org/10.4321/S2014-98322014000400005

Schön, D. A. (1992). The theory of inquiry: Dewey's legacy to education. *Curriculum Inquiry, 22*(2), 119-139.

Schwarzer, R., & Jerusalem, M. (1993). *Measurement of Perceived Self-Efficacy: Psychometric Scales for Cross-Cultural Research*. Freie University.

Turner, S., & Harder, N. (2018). Psychological safe environment: A concept analisis. *Clinical Simulation in Nursing, 18*(5), 47-55. https://doi.org/10.1016/j.ecns.2018.02.004