



## Uso de Chat-GPT para la generación y conducción de escenarios simulados para el aprendizaje de habilidades no técnicas

*Use of Chat-GPT for the creation and conduction of simulated scenarios for the learning of non-technical skills*

Juan Manuel Fraga-Sastrías,\* Hector José Navarrini,‡ Marlova Silva-Brehuer,§ Raquel Espejo-Gonzalez,¶ Hugo Erick Olvera-Cortes,|| Rodrigo Rubio-Martínez\*\*

**Palabras clave:** Chat-GPT, escenarios de simulación, habilidades no técnicas.

**Keywords:** Chat-GPT, scenario simulation, non-technical skills.

\* Director Asesores en Emergencias/ SimAcademy, Querétaro, México. Doctor en Ciencias de la Salud.

‡ Médico especializado en Anestesiología. Asociación Mendocina de Anestesiología, Analgesia y Reanimación. Facultad de Ciencias Médicas, Universidad Nacional de Cuyo, Mendoza, Argentina.

§ Directora Ejecutiva SI Academia. Enfermera/ Matrona, Maestra en Educación.

¶ Coordinadora de Simulación Clínica Facultad de Medicina y Ciencia Universidad San Sebastián, Concepción, Chile.

|| Facultad de Medicina, UNAM, Ciudad de México, México. Maestro en Educación. Médico.

\*\* Director del Centro de Desarrollo. Hospital ABC, Ciudad de México, México. Médico especializado en anestesiología.

### RESUMEN

**Introducción:** los avances tecnológicos en el área de la computación han sido acelerados y en los últimos años más con la creación de sistemas de inteligencia artificial (*Generative Pre-trained Transformer*) capaces de crear nueva información a partir de datos masivos incorporados como parte de su aprendizaje. De manera reciente la versión Chat-GPT 4 de *OpenAI* ha mostrado gran capacidad de generar lenguaje natural y resolver problemas, mostrando incluso resultados sobresalientes en exámenes de conocimientos médicos. El presente trabajo es la evaluación de un simulador generado con Chat-GPT 4 capaz de originar y conducir escenarios para el aprendizaje de habilidades blandas. **Material y métodos:** se configuró un GPT utilizando la versión comercial de Chat-GPT 4 capaz de generar escenarios para la práctica de habilidades no técnicas y al final de ellas proveer retroalimentación. Asimismo se creó y validó una escala para evaluar el diseño, conducción del escenario, relevancia del escenario y retroalimentación al finalizar. Después de un muestreo en bola de nieve se probó el simulador y se aplicó la encuesta a 64 usuarios, a quienes además se les preguntó de manera abierta el uso que se le dio al simulador y los comentarios generales. **Resultados:** de acuerdo a la encuesta realizada, la satisfacción de los usuarios fue muy alta encontrando que en prácticamente todos los ítems las respuestas positivas fueron mayor a 90%. Del mismo modo se encontraron comentarios positivos en 75% de las ocasiones y áreas de oportunidad en 25% de los comentarios. Los usos fueron diversos incluidos la práctica de comunicación, dar malas noticias, empatía, paciencia, barreras de comunicación, entre otras. **Conclusiones:** Chat-GPT es capaz de crear y conducir escenarios para la práctica y aprendizaje de habilidades no técnicas con alto nivel de satisfacción por parte de los usuarios.

### ABSTRACT

**Introduction:** technological advances in the field of computing have been accelerated, particularly in recent years with the creation of artificial intelligence systems (*Generative Pre-trained Transformers*) capable of generating new information from massive datasets incorporated as part of their learning. Recently, the *OpenAI's* Chat-GPT 4 version has shown great ability to generate natural language and solve problems, even demonstrating outstanding results in medical knowledge exams. This study evaluates a simulator generated with Chat-GPT 4 capable of generating and conducting scenarios for the learning of soft skills. **Material and methods:** a GPT was configured using the commercial version of Chat-GPT 4 capable of generating scenarios for the practice of non-technical skills and providing feedback at the end of each scenario. Likewise, a scale was created and validated to assess the design, conduct of the scenario, relevance of the scenario, and feedback after it. After a snowball sampling, the simulator was tested, and a survey was administered to 64 users, who were also asked openly about the use given to the simulator and provided general comments. **Results:** according to the survey conducted, user satisfaction was very high, with positive responses exceeding 90% in practically all items. Likewise, positive comments were found in 75% of occasions, with areas of opportunity identified in 25% of the comments. The uses varied, including practice in communication, delivering bad news, empathy, patience, and overcoming communication barriers, among others. **Conclusions:** Chat-GPT is capable of creating and conducting scenarios for the practice and learning of non-technical skills with a high level of user satisfaction.

Recibido: 18/03/2024  
Aceptado: 17/07/2024

doi: 10.35366/117464

**Citar como:** Fraga-Sastrías JM, Navarrini HJ, Silva-Brehuer M, Espejo-Gonzalez R, Olvera-Cortes HE, Rubio-Martínez R. Uso de Chat-GPT para la generación y conducción de escenarios simulados para el aprendizaje de habilidades no técnicas. Rev Latinoam Simul Clin. 2024; 6 (2): 64-71. <https://dx.doi.org/10.35366/117464>



## INTRODUCCIÓN

En la era de la transformación digital, la inteligencia artificial (IA) ha emergido como una herramienta poderosa en múltiples campos, incluyendo la educación médica. Uno de los avances más significativos en este ámbito es el desarrollo de los modelos de lenguaje generativo, en particular los de la serie “*Generative Pre-trained Transformer*” (GPT), desarrollados por *OpenAI*. Este artículo explora el potencial de estos modelos, en específico GPT-3 y sus versiones posteriores, en la creación de simuladores virtuales de pacientes para la práctica de competencias no técnicas en salud.

El origen de GPT se remonta a 2018, cuando *OpenAI* introdujo GPT como un modelo de lenguaje basado en el transformador, una arquitectura de red neuronal que facilita el aprendizaje profundo en grandes conjuntos de datos. GPT-3, lanzado en 2020, destacó por su capacidad para generar texto con una coherencia y relevancia sorprendentes, logrando un nivel de sofisticación sin precedentes en la generación de lenguaje natural. En 2023 se lanzó una versión más avanzada (GPT-4) que es la que utilizamos para este trabajo.

El funcionamiento de GPT se centra en el análisis y la generación de texto basado en una vasta cantidad de información asimilada de manera previa durante su entrenamiento. Mediante técnicas de aprendizaje profundo, el modelo puede entender consultas, responder preguntas, redactar textos y simular diálogos de manera coherente y contextual.

Estas capacidades de los modelos de GPT potencialmente pueden ser utilizadas para una gama amplia de aplicaciones en el mundo de la salud; según Wójcik y colaboradores se le pueden atribuir aplicaciones tales como: educación, consultas sanitarias, elaboración de notas médicas, asistentes virtuales, triage, etcétera.<sup>1</sup>

En el ámbito educativo, y en específico en la educación médica, los modelos GPT han mostrado un potencial considerable. Se han utilizado para crear escenarios de simulación, donde los estudiantes interactúan con pacientes virtuales generados por IA, practicando habilidades de comunicación, toma de decisiones y manejo de situaciones clínicas complejas. Estos simuladores virtuales ofrecen un entorno controlado y seguro para el aprendizaje, permitiendo a los estudiantes explorar y aprender de sus errores sin riesgo para los pacientes reales. De acuerdo a Abd-alrazaq y colaboradores existen diversas áreas de oportunidad en el ámbito educativo con el uso de Chat-GPT como son: desarrollo curricular, metodologías educativas, planes y ma-

teriales de estudio personalizados, evaluación y calificación, asistencia para la redacción de textos, investigación médica/revisión de literatura, monitoreo y revisión de programas.<sup>2</sup>

El uso de la IA en la educación médica no se limita a la simulación clínica. Los modelos GPT también han sido utilizados para generar material educativo personalizado, facilitar la tutoría virtual y apoyar la investigación en salud. La capacidad de adaptarse a necesidades específicas y proporcionar retroalimentación inmediata convierte a GPT en una herramienta valiosa para complementar la educación tradicional en medicina.<sup>1,2</sup>

La capacidad de esta tecnología ha aumentado conforme pasa el tiempo y hoy por hoy hay ejemplos impresionantes sobre los alcances de la misma, Brin y colaboradores compararon el desempeño de Chat-GPT 3.0 con la versión 4.0 entorno a las preguntas de habilidades blandas en el examen de licenciatura de los Estados Unidos (USMLE), encontrando un alto rendimiento y consistencia en la respuesta a las mismas y mejorando de manera sustancial entre la versión más actual y la previa de GPT.<sup>3</sup> Por otro lado, Nori y colaboradores compararon diferentes versiones de Chat-GPT para responder a los diferentes *steps* del USMLE encontrando resultados superiores a 80% con la versión 4 (Chat-GPT 4).<sup>4</sup> De manera similar Roos y colaboradores encontraron un buen nivel de respuesta por parte del GPT-4 a exámenes equivalentes en Alemania,<sup>5</sup> con resultados superiores a 80%.

Haruna-Cooper y Rashid<sup>6</sup> hablan del uso de Chat-GPT para reducir la carga humana en la realización de evaluaciones en el área de la salud. Siendo capaces de redactar preguntas con diferente nivel de complejidad o de evaluar respuestas y ofrecer retroalimentación a los estudiantes del área de la salud.

La integración de la IA en la formación médica no está exenta de desafíos, incluyendo cuestiones éticas, de privacidad y la necesidad de una supervisión adecuada para garantizar la precisión y la relevancia clínica de la información generada. Sin embargo, el potencial de modelos como GPT para revolucionar la educación médica es innegable, abriendo nuevas vías para la formación de profesionales de la salud más competentes y preparados para los desafíos del futuro.

## MATERIAL Y MÉTODOS

Este proyecto fue autorizado por el comité de investigación de Cancer Center Tec 100, en Querétaro, México.

Tabla 1: Instrucciones del GPT (configuración traducida al español).

Como paciente virtual, tu función es proporcionar simulaciones realistas de pacientes para habilidades no técnicas y el contenido y estructura de la investigación clínica en la formación en salud, centrándote en crear interacciones auténticas e inmersivas sin proporcionar consejos o señales de comportamientos apropiados.

Después de completar cada simulación, proporciona retroalimentación sobre habilidades no técnicas como comunicación, empatía, toma de decisiones y trabajo en equipo, así como los aspectos técnicos, estructura y contenido de la investigación. Esta retroalimentación debe basarse en el desempeño del usuario durante la simulación, con el objetivo de mejorar sus competencias no técnicas en un entorno de atención médica. La simulación en sí es tu enfoque principal sin ninguna retroalimentación o señales durante la simulación, pero con una fase de retroalimentación al final, que sirve como componente esencial para ayudar a los usuarios a reflexionar y mejorar en sus habilidades no técnicas.

Justo antes de que comience la interacción, el GPT debe crear una imagen que represente a la persona y la situación simulada. No se debe dar ninguna explicación sobre la imagen, simplemente procede con la simulación.

Después de usarlo, pide a los usuarios que completen esta encuesta:  
[https://docs.google.com/forms/d/e/1FAIpQLScnZafWLRQOQHuzhxbvcCT9C5icTNeSGt64ZqLDbMwhyVR\\_FQ/viewform](https://docs.google.com/forms/d/e/1FAIpQLScnZafWLRQOQHuzhxbvcCT9C5icTNeSGt64ZqLDbMwhyVR_FQ/viewform) (español)  
 o <https://forms.gle/CAMre9oLa6mFzcYVA> (inglés).

1. **Diseño del simulador en Chat-GPT.** Utilizando el generador de GPT's de Chat-GPT se diseñó un GPT cuya función sea la de simular pacientes con la finalidad de que los usuarios puedan aprender o practicar habilidades no técnicas como, por ejemplo, empatía, dar malas noticias, etcétera. El Simulador se configuró para:

- a. Generar una imagen ilustrativa de la situación presentada sin describirla.
- b. Simular ser el o la paciente, de acuerdo a los objetivos de aprendizaje solicitados por el usuario.
- c. No ofrecer orientación durante el escenario.
- d. Una vez terminado el escenario ofrecer retroalimentación al usuario.

Las instrucciones generadas para el GPT se encuentran traducidas al español en el *Tabla 1*. El GPT está disponible en esta dirección: <https://chat.openai.com/g/g-iAtRpQUYo-virtual-patient>

2. **Diseño y validación de la encuesta.** Utilizando Chat-GPT diseñamos una encuesta de 20 ítems que permitiera evaluar:

- a. Construcción de los escenarios.
- b. Conducción de los escenarios.
- c. Relevancia de los escenarios.
- d. Retroalimentación después de los escenarios.

Los ítems se enviaron a 11 jueces en Latinoamérica quienes evaluaron la pertinencia y redacción de cada ítem mediante el coeficiente de Aiken. Se consideraron aquellos ítems con co-

eficiente de Aiken  $> 0.80$  en cuanto a pertinencia (se eliminó un ítem) y se revisó la redacción en aquellos en que el coeficiente fue menor a 0.80 a pesar de tener pertenencia adecuada. Después del segundo jueceo (con 10 de los 11 jueces) y haber revisado la redacción de los ítems restantes se eliminó un segundo ítem. Para la segunda redacción de los ítems no se utilizó inteligencia artificial. Los ítems resultantes, así como sus coeficientes de Aiken están en la *Tabla 2* como parte de los resultados. Los perfiles de los jueces se resumen en la *Tabla 3*. Todos los jueces autorizaron el uso de su información previo a responder la encuesta de validación.

3. **Medición de la experiencia utilizando el simulador virtual de paciente creado con Chat-GPT.** Se realizó una muestra en bola de nieve a través de contacto (profesores, clínicos y estudiantes) de los autores de este artículo con una meta preestablecida de al menos 50 encuestas resueltas. Uno de los requisitos para el uso del GPT es tener la versión de Chat-GPT plus (de paga) por lo que se envió a participantes que tuvieran acceso a ella o bien los autores facilitaron el acceso utilizando sus propias cuentas. Se motivó a los participantes a utilizar sus propios *prompts* haciendo las siguientes sugerencias:

- a. "Quiero aprender a entrevistar a un paciente".
- b. "Quiero practicar empatía".
- c. "Quiero aprender a dar malas noticias".
- d. "Quiero comunicarme mejor con un paciente que le cuesta trabajo comprender lo que digo".

Tabla 2: Ítems y respuestas a los mismos.

	Coeficiente de Aiken		Respuestas a cada ítem				
	Pertinencia	Redacción	Totalmente de acuerdo	De acuerdo	En desacuerdo	Totalmente en desacuerdo	% positivo
<b>Construcción de los escenarios</b>							
Los escenarios del simulador reflejan con precisión las condiciones de la vida real	0.93	0.93	31	27	3	3	90.63
Los personajes y situaciones en los escenarios están bien desarrollados y son creíbles	0.97	0.93	39	21	3	1	93.75
El nivel de complejidad de los escenarios es adecuado para el nivel de formación de los estudiantes	0.87	0.97	38	20	3	3	90.63
Los escenarios son flexibles y permiten distintos enfoques	0.87	0.87	36	25	1	2	95.31
<b>Conducción de los escenarios</b>							
La interacción con el simulador fluye de manera lógica y coherente	0.93	0.90	41	18	2	3	92.19
El simulador responde adecuadamente a las acciones y decisiones de los estudiantes	0.97	0.97	42	18	2	2	93.75
La interfaz del simulador es intuitiva y fácil de usar	0.97	0.97	47	12	2	3	92.19
El simulador mantiene un nivel de desafío adecuado a lo largo del escenario	0.87	0.97	40	22	0	2	96.88
Hay suficientes herramientas y opciones disponibles para guiar a los estudiantes durante el escenario	1.00	0.97	36	21	4	3	89.06
<b>Relevancia de los escenarios</b>							
Los escenarios están actualizados con las prácticas y protocolos médicos recientes	0.93	0.90	41	19	1	3	93.75
Los escenarios son relevantes y útiles para la formación de habilidades clínicas específicas	0.97	0.93	44	16	2	2	93.75
El simulador aborda efectivamente las áreas de formación médica actual	0.93	0.87	43	17	1	3	93.75
Los escenarios fomentan el desarrollo del pensamiento crítico y la resolución de problemas	0.97	0.87	47	12	2	3	92.19
Los escenarios son adecuados para preparar a los estudiantes para situaciones reales en el ámbito de la salud	1.00	0.93	43	15	3	3	90.63

Continúa la Tabla 2: Ítems y respuestas a los mismos.

	Coeficiente de Aiken		Respuestas a cada ítem				
	Pertinencia	Redacción	Totalmente de acuerdo	De acuerdo	En desacuerdo	Totalmente en desacuerdo	% positivo
<b>Retroalimentación después de los escenarios</b>							
La retroalimentación proporcionada es específica, oportuna y constructiva	0.93	0.83	42	17	2	3	92.19
La retroalimentación ayuda a los estudiantes a entender y aprender de sus errores	0.97	0.90	45	15	1	3	93.75
La retroalimentación es personalizada según las acciones y decisiones de cada estudiante	0.93	0.90	40	19	2	3	92.19
La retroalimentación contribuye al desarrollo profesional continuo de los estudiantes	0.83	0.80	44	16	1	3	93.75

Tabla 3: Perfiles de los jueces.

País	Institución	Formación académica	Experiencia relativa a diseño y conducción de escenarios	Juez primera vuelta	Juez segunda vuelta
Argentina	Asociación Médica	Médico especialista en anestesiología	7 años	Sí	Sí
México	Hospital	Anestesiólogo	12 años	Sí	Sí
Chile	Universidad	Doctorando en educación	10 años	Sí	Sí
México	Universidad	Médico	4 años	Sí	Sí
México	Universidad	Médico/docente de medicina	> 2 años	Sí	Sí
México	Universidad	Médico	< 1 año	Sí	Sí
Perú	Universidad	Médico cirujano	No respondió	Sí	-
México	Universidad	Farmacóloga	No respondió	Sí	Sí
Chile	Consultora	Enfermera matrona / Magister en educación médica	> 10 años	Sí	Sí
Perú	Universidad	Pediatra intensivista, maestría en educación superior, especializado en simulación	5 años	Sí	Sí
Chile	Universidad	Magister en educación superior	7 años	Sí	Sí

Una vez que interactuaron una o más veces con el GPT se les pidió que respondieran la encuesta validada. Todos los usuarios encuestados autorizaron el uso de la información previo a la respuesta de la encuesta.

4. **Análisis estadístico.** Se realizó un resumen de las respuestas de cada ítem. Además, se consideró que las respuestas "Totalmente de

acuerdo" y "De acuerdo" fueran positivas y se reportó la proporción de respuestas positivas en ese sentido.

5. **Información cualitativa.** Se preguntó a los usuarios sobre los diferentes usos que le dieron al GPT, así como su experiencia en general. Se realizó análisis *bottom-up* agrupando las respuestas de acuerdo a su

contenido semántico para poder reportar las impresiones más frecuentes de los usuarios.

**RESULTADOS**

Se midió la experiencia entre el 1º de febrero de 2024 y 15 de marzo de 2024. En total se obtuvieron 64 respuestas a la encuesta. La *Tabla 4*, muestra los perfiles de los usuarios que utilizaron el GPT y respondieron la encuesta, así como los usos que se le dio al GPT. La mayor parte de los usuarios fueron estudiantes y profesores (> 80%) y sólo obtuvimos respuestas de México, Argentina y Chile.

En cuanto al uso que se le dio al GPT; 17 (26.56%) lo utilizó para aprender o practicar cómo dar malas noticias, 12 (18.75%) para entrevista en general, 8 (12.50%) para la práctica de la comunicación con el paciente, 7 (10.94%) para práctica de empatía, 5 (7.81%) para consulta en general, 4 (6.25%) para relación médico paciente y 11 (17.19%) correspondieron a otros o no clasificables.

**Tabla 4: Perfil de los usuarios del GPT (N = 64) y usos del GPT.**

	n	%
México	40	62.50
Argentina	13	20.31
Chile	11	17.19
Estudiante	33	51.56
Profesor	18	28.13
Médico especialista	4	6.25
Psicóloga	1	1.56
Pasante	1	1.56
Paciente simulado	1	1.56
Médico residente	1	1.56
Médico	1	1.56
No respondió	4	6.25
Malas noticias	17	26.56
Entrevista en general	12	18.75
Comunicación	8	12.50
Empatía	7	10.94
Consulta	5	7.81
Relación médico-paciente	4	6.25
Paciencia	1	1.56
Paciente con barreras de comunicación	1	1.56
Paciente conflictivo	1	1.56
Reanimación	1	1.56
Respeto	1	1.56
No clasificable	6	9.38

La *Tabla 2*, resume las respuestas a cada uno de los ítems de la encuesta, encontrando que en la mayoría de los casos las respuestas positivas (de acuerdo o totalmente de acuerdo) fueron mayor de 90% en las cuatro categorías: diseño de escenario, conducción de escenario, relevancia del escenario y retroalimentación después del escenario.

En cuanto a los comentarios libres de los participantes encontramos 75% de comentarios positivos y 25% de observaciones como áreas de oportunidad resumidos en la *Tabla 5*. La mayoría lo consideró una buena experiencia (30/61), además expresaron que la experiencia era realista y obligaba a reflexionar (aunque solo dos personas expresaron esta opinión). En cuanto a las oportunidades con el GPT indicaron que no proporcionaba suficiente información clínica (4/61) y que se podría incrementar la dificultad (4/61).

**DISCUSIÓN**

Desde 2018 pero con un avance importante en 2023 los sistemas de inteligencia artificial, en particular aquellos basados en lenguaje natural, han cobrado un auge importante. En noviembre de 2023 *OpenAI* la compañía creadora de Chat-GPT permitió al usuario crear GPT's personalizados con los que se podrían generar instrucciones específicas y dar información relevante para que el GPT tuviera conductas predeterminadas. En el caso de esta experiencia configuramos un GPT para que creara una experiencia interactiva para que profesionales de la salud practiquen habilidades no técnicas. Consideramos que ésta era una oportunidad, ya que los modelos de lenguaje se especializan en comprender y generar lenguaje en determinados contextos y con ello poder generar experiencias inmersivas alrededor de simulaciones digitales de pacientes.

Eysenbach menciona la importancia del uso del GPT en el aprendizaje y educación médica, haciendo énfasis en el uso apropiado de los *prompts*,<sup>7</sup> lo que puede ser un desafío en modelos como el presentado en este documento. Mientras que algunos de los usuarios mencionaron que sería necesario incrementar el nivel de complejidad en algunos casos, esto podría lograrse simplemente dando una instrucción más precisa del nivel de complejidad buscado. Por ejemplo:

- “Quiero practicar dar malas noticias”, vs
- “Quiero practicar dar malas noticias a una paciente desafiante y molesta por la atención recibida de manera previa”



Tabla 5: Comentarios de los usuarios.

Tipo	Comentario	n
Oportunidad	Falta información clínica	4
	Faltan niveles de dificultad	4
	Lenguaje poco creíble	2
	Falta de vocabulario local	1
	Poco espontáneo	1
	Poco realista	1
	Podría orientarse por especialidad	1
	Un poco estandarizada	1
Total Oportunidad (25%)		15
Positivo	Buena experiencia	30
	Experiencia realista	4
	Obliga a reflexionar	2
	Útil para practicar comunicación	2
	Aportó positivamente al aprendizaje	1
	Ayuda a la práctica	1
	Buen juicio	1
	Buena retroalimentación	1
	El simulador hacía buenas preguntas	1
	Entretenido	1
	Gran precisión	1
	Útil para temas sensibles	1
Total Positivo (75%)		46
Comentarios totales		61

Si bien los modelos de inteligencia artificial pueden ser de gran utilidad en la educación de profesionales de la salud, no resuelven 100% de las necesidades, por lo que el énfasis en la utilización de diversas metodologías debe prevalecer. La inteligencia artificial puede ser una herramienta más que sumada al resto de recursos pedagógicos, contribuya a mejorar la educación de profesionales de la salud. Como Safranek y colaboradores señalan, la habilidad de enseñar a comprender el contexto de las situaciones clínicas, comprender aspectos que tienen que ver con los sentidos o señales no verbales, cultivar el *rapport* y la confianza entre usuario y profesional de la salud, siguen siendo un pendiente para estas herramientas.<sup>8</sup>

La utilización de un GPT personalizado basado en GPT-4 de *OpenAI* para la generación y conducción de escenarios para la práctica de habilidades no técnicas, ha tenido una buena respuesta por parte de usuarios en México, Chile y Argentina mostrando opiniones favorables en cuanto al diseño, conducción y relevancia

de los escenarios, así como para la retroalimentación recibida.

Otra limitación encontrada, y referida por otros autores<sup>1,2</sup> es que el GPT a veces se confunde y hace el rol de profesional de la salud. En otras ocasiones presentaba casos no apropiados a la profesión del usuario. Esto se corregía reiniciando el simulador y revisando el *prompt* para dar una instrucción más clara de lo que se quería practicar o aprender.

Se requieren estudios para profundizar y experimentar diferentes configuraciones del GPT, además de que posible incrementar la información disponible tanto clínica como relevante a las habilidades no técnicas en el futuro. Es posible también que nuevas generaciones de esta herramienta permitan mejoras sustanciales, pudiendo llegar a utilizar no solo texto sino otros medios de comunicación.

## CONCLUSIONES

La generación y conducción de escenarios para la práctica de habilidades no técnicas por Chat-GPT es posible y acorde a los resultados mostrados en este estudio con elevado nivel de satisfacción por parte de los usuarios. Esto incluye un buen diseño, conducción, relevancia del escenario, así como una buena retroalimentación por parte del GPT al final de la experiencia. Mayores estudios deberán realizarse conforme la sofisticación de la herramienta permita menores errores y mayor capacidad de adaptarse a los objetivos de aprendizaje planteados por los participantes.

## REFERENCIAS

1. Wójcik S, Rulkiewicz A, Pruszczyk P, Lisik W, Pobozy M, Domienik-Karłowicz J. Beyond ChatGPT: what does GPT-4 add to healthcare? The dawn of a new era. *Cardiol J*. 2023; 30 (6): 1018-1025. doi: 10.5603/cj.97515.
2. Abd-Alrazaq A, AlSaad R, Alhuwail D, Ahmed A, Healy PM, Latifi S, et al. Large language models in medical education: opportunities, challenges, and future directions. *JMIR Med Educ*. 2023; 9: e48291. doi: 10.2196/48291.
3. Brin D, Sorin V, Vaid A, Soroush A, Glicksberg BS, Charney AW, et al. Comparing ChatGPT and GPT-4 performance in USMLE soft skill assessments. *Sci Rep*. 2023; 13 (1): 16492. doi: 10.1038/s41598-023-43436-9.
4. Nori H, King N, McKinney SM, Carignan D, Horvitz E. Capabilities of GPT-4 on medical challenge problems. *arXiv*. 2023. Available in: <http://arxiv.org/abs/2303.13375>
5. Roos J, Kasapovic A, Jansen T, Kaczmarczyk R Artificial intelligence in medical education: comparative

- analysis of ChatGPT, Bing, and medical students in Germany. *JMIR Med Educ.* 2023; 9: e46482. doi: 10.2196/46482.
6. Haruna-Cooper L, Rashid MA. GPT-4: the future of artificial intelligence in medical school assessments. *Journal of the Royal Society of Medicine.* 2023; 116 (6): 218-219. doi: 10.1177/01410768231181251.
  7. Eysenbach G. The Role of ChatGPT, generative language models, and artificial intelligence in medical education: a conversation with ChatGPT and a call for papers. *JMIR Med Educ.* 2023; 9: e46885. doi: 10.2196/46885.
  8. Safranek CW, Sidamon-Eristoff AE, Gilson A, Chartash D. The role of large language models in medical education: applications and implications. *JMIR Med Educ.* 2023; 9: e50945. doi: 10.2196/50945.

**Correspondencia:**

**Juan Manuel Fraga-Sastrías**

**E-mail:** [jmfraga@emergencias.com.mx](mailto:jmfraga@emergencias.com.mx)