



ORIGINAL ARTICLE

Experiencia de uso de una herramienta para evaluar el realismo en simulaciones clínicas



Judit Sánchez Gómez^{a,*}, Gleyvis Coro-Montanet^b, Helena Wagner Porto Rocha^d,
Esther García García^d y Óscar Oliva Fernández^c

^a Facultad de Medicina, Deporte y Salud, Universidad Europea de Madrid, Madrid, España

^b Facultad de Ciencias de la Vida y la Naturaleza, Universidad de Nebrija, Madrid, España

^c Departamento Hospital Simulado, Facultad de Medicina, Deporte y Salud, Universidad Europea de Madrid, Madrid, España

^d Facultad de Ciencias Biomédicas y de la Salud, Universidad Europea de Madrid, Madrid, España

Recibido el 6 de mayo de 2024; aceptado el 11 de julio de 2024

Disponible en Internet el 24 de octubre de 2024

PALABRAS CLAVE

ProRealSim;
Simulación;
Estudio cualitativo;
Evaluación;
Realismo;
Evaluación de realismo

Resumen

Introducción: la medición del realismo en la simulación clínica es un desafío complejo debido a su naturaleza subjetiva y multidimensional. ProRealSim es una herramienta de medición de realismo desarrollada por investigadores de la Universidad Europea de Madrid, basada en un marco teórico que categorizó el realismo en 3 dimensiones: participante simulado, escenografía y simulador.

Material y métodos: el objetivo de este estudio fue conocer la percepción que tiene un equipo de expertos clínicos sobre la herramienta ProRealSim. Para ello, se realizó una entrevista semiestructurada a un grupo focal, compuesto por 4 expertos en el campo de la reanimación cardiopulmonar (RCP).

Resultados: Los resultados del estudio mostraron una percepción positiva por parte de los expertos hacia la herramienta ProRealSim, resaltando su intuitivo diseño y la claridad de los indicadores presentados. No obstante, también se identificaron áreas susceptibles de mejora, siendo la adaptación para dispositivos móviles una de las principales sugerencias. En lo referente al contenido del medidor, los especialistas evaluaron los indicadores como pertinentes y de fácil comprensión. No obstante, se señaló la necesidad de revisar la relevancia de ciertos indicadores que podrían no aplicar en todas las simulaciones, así como la subjetividad inherente en la evaluación, proponiendo la participación de múltiples tasadores para garantizar una evaluación más objetiva.

Conclusión: en general, los hallazgos del estudio ofrecieron una valiosa perspectiva para la realización de evaluaciones utilizando la herramienta ProRealSim. Ante la complejidad de evaluar el realismo en simulaciones clínicas, se sugieren estrategias tales como capacitación

* Autor para correspondencia.

Correo electrónico: judit.sanchez@universidadeuropea.es (J. Sánchez Gómez).

previa y consenso entre los tasadores para mejorar la efectividad y objetividad de las mediciones.

© 2024 Los Autores. Publicado por Elsevier España, S.L.U. Este es un artículo Open Access bajo la licencia CC BY-NC (<http://creativecommons.org/licenses/by-nc/4.0/>).

KEYWORDS

ProRealSim;
Simulation;
Qualitative study;
Evaluation;
Realism;
Realism evaluation

Experience of using a tool to assess realism in clinical simulations

Abstract

Introduction: Measuring realism in clinical simulation poses a complex challenge due to its subjective and multidimensional nature. ProRealSim, a realism measurement tool developed by researchers from Universidad Europea de Madrid, is based on a theoretical framework that categorizes realism into three dimensions: simulated participant, scenography, and simulator.

Materials and methods: The study aimed to comprehend the perception of clinical experts regarding the ProRealSim tool. A focus group, comprising four clinical experts (three nurses and one doctor) with extensive experience in cardiopulmonary resuscitation (CPR) simulations, was conducted through a semi-structured interview.

Results: The study revealed that the experts held a positive view of the ProRealSim website, emphasizing its user-friendliness and the clarity of the indicators. However, they also identified potential improvements, such as optimizing the version for mobile devices. Regarding the tool's content, experts found the indicators suitable and easily understandable. Nevertheless, they highlighted certain aspects for improvement, including the relevance of certain indicators that may not apply to all simulations and the inherent subjectivity of the evaluation, recommending the involvement of more than one evaluator in the assessment.

Conclusions: Overall, the study results provide valuable insights for conducting assessments with the ProRealSim tool. The inherent complexity in assessing realism in simulations was highlighted, and strategies such as pre-training, identifying relevant variables for each simulation, and objectivity in such assessments were suggested.

© 2024 The Authors. Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>).

Introducción

En el aprendizaje basado en simulación, el realismo es una variable tan fundamental que ha servido para clasificar los tipos de escenarios que se realizan en simulaciones de baja, media o alta fidelidad¹. No obstante, al examinar la literatura, se constata que la mayoría de estas clasificaciones parten de percepciones subjetivas. De este modo, son escasas y bastante limitadas las herramientas de medición que han sido validadas estadísticamente y aplicadas en muestras amplias, dificultándose la interpretación objetiva y científica del realismo¹⁻⁵.

Algunos autores han insistido en la necesidad de categorizar y subcategorizar el realismo para poder interpretarlo de una manera más objetiva y certera⁶⁻⁸. En respuesta a este vacío conceptual y evaluativo, un equipo de investigadores de la Universidad Europea de Madrid desarrolló un nuevo marco teórico y un instrumento de medición cuantitativa llamado ProRealSim⁹. La herramienta se digitalizó en una página web con 2 pestañas: un home o página principal, donde se explicitaron las categorías del realismo a manejar para la evaluación, y una página con el medidor propiamente dicho, con todos los ítems a valorar. Entre sus ventajas sobresalieron la utilización de indicadores objetivos, medibles, fáciles de interpretar y familiares a la

mayoría de los escenarios simulados. Esto posibilitó la obtención de un índice de realismo calculado mediante fórmulas validadas estadísticamente¹⁰, cuyo principal usuario fue el docente diseñador experto, favorecido por la amplia disponibilidad digital y carácter trilingüe (español, inglés, portugués) de la herramienta que permitió hacer estudios con grandes muestras de múltiples contextos y facilitó la comparación entre mediciones expertas de realismo conseguido, con las percepciones del realismo reportadas por el alumnado o participantes. Este enfoque ayudó a reducir la consideración, casi siempre subjetiva, del fenómeno simulado, superar el disenso y la falta de herramientas evaluativas, y prever y obtener el verdadero coste/beneficio de la inversión en entornos realistas e hiperrealistas. De igual forma, tener una medida objetiva del realismo facilitará valorar científicamente el impacto del mismo en el aprendizaje, la psicología y la bioética de los estudiantes que se forman con esta metodología¹¹.

La herramienta ProRealSim entrañó cierto grado de dificultad en su utilización al tratarse de un instrumento de evaluación novedoso del que no se poseía una experiencia sistemática de uso^{8,12}. Con independencia de su necesidad y utilidad, la falta de una cultura de medición e interpretación de esta variable influyó en la aplicación de la herramienta. Del mismo modo, cuando un evaluador no tiene experiencia

previa resultan frecuentes los errores de apreciación e interpretación.

Con base en ello, se realizó un análisis exhaustivo con el objetivo de conocer la percepción que tuvieron los expertos clínicos sobre el instrumento de medición ProRealSim.

Materiales y métodos

La metodología empleada en este estudio es de carácter cualitativo¹³, basada en el paradigma interpretativo^{14,15}, que permitió comprender la subjetividad de los expertos dentro del contexto evaluativo y su experiencia.

Para llevar a cabo el estudio, se realizó una entrevista semiestructurada con un grupo focal compuesto por 4 evaluadores expertos en el campo de la rehabilitación cardiopulmonar (RCP). Previamente a la entrevista, estos evaluadores habían realizado las mediciones utilizando ProRealSim en el hospital simulado de la Universidad Europea de Madrid. Las simulaciones se centraron específicamente en escenarios de RCP.

Muestreo y reclutamiento de los participantes

Dada la necesidad de entrevistar a aquellos expertos más familiarizados con la herramienta y poder conducir el grupo focal en la búsqueda de datos ricos y fiables sobre su funcionamiento, se realizó un muestro por conveniencia¹⁶.

El grupo focal estuvo compuesto por 2 enfermeros y 2 médicos (un hombre y una mujer, en ambos casos). Todos los expertos en RCP fueron previamente capacitados en el manejo de la herramienta ProRealSim, completando un curso online de una hora sobre el uso de la herramienta en su entorno web. Los 2 enfermeros habían aplicado la herramienta en más de 20 escenarios de enfermería. Uno de los médicos la había aplicado en 10 escenarios y otro la aplicaba por primera vez. Aunque uno de los médicos tenía menos experiencia directa con la herramienta ProRealSim, su experiencia general en el ámbito clínico en RCP y su formación con la herramienta, justifica su inclusión como experto en el grupo focal.

Los datos fueron recolectados en diciembre de 2022. De los 4 expertos, 3 eran docentes en la Universidad Europea de Madrid y el médico con menor experiencia era anestesista en el servicio de urgencias de la Comunidad de Madrid.

Procedimiento

El grupo focal fue conducido por una experta en el manejo de la herramienta ProRealSim. Aunque no tenía experiencia en realizar entrevistas ni formó parte de las mediciones de RCP, su conocimiento sobre la herramienta fue esencial para conducir la entrevista correctamente. Se invitó a los participantes a describir sus experiencias con la herramienta y dar sus puntos de vista sobre aspectos positivos y/o negativos. Se partió de preguntas generales, derivando hacia preguntas más específicas¹⁷. La conducción del grupo fue semiestructurada, para poder abordar los temas principales de interés para la investigación, aunque no se cerró a la posible aparición de nuevos temas¹⁸. La sesión se condujo de manera que se trataron los diferentes temas hasta llegar a la saturación de los datos.

El consentimiento se obtuvo por escrito antes de la entrevista y también de manera verbal al comenzar la misma. La entrevista tuvo una duración de una hora y 56 minutos, fue grabada en vídeo y se descargó la transcripción con Microsoft Teams para su posterior análisis.

Análisis de los datos

Una vez terminada la discusión del grupo focal, se revisó la transcripción y se corrigieron posibles errores. Posteriormente, se realizó análisis temático deductivo¹⁹. Se partió de 2 temas principales: el funcionamiento de la herramienta (relacionado con el análisis de las características funcionales de las pestañas de la web que alojaba a la herramienta en su relación con el usuario evaluador) y el contenido del medidor (relacionado con el análisis del marco evaluativo del medidor, indicaciones para el cálculo y pertinencia para evaluar los niveles de realismo en las diferentes dimensiones y unidades de análisis expresadas a través de los indicadores). El procedimiento que se siguió para el análisis cualitativo fue: lectura de los datos, categorización de los verbatim, inclusión en subtemas y de estos, a su vez, en códigos, para posteriormente realizar la descripción de los resultados²⁰. Después, se interpretaron los códigos con base en los verbatim de los participantes y se establecieron conexiones entre ellos²¹. El análisis se realizó mediante la herramienta Excel.

Para aumentar la confiabilidad de los datos se realizó una nueva revisión del análisis por parte de la conductora de la entrevista y de una observadora, también experta en el manejo de ProRealSim¹⁶.

Resultados

En esta sección, se presentan los resultados del análisis cualitativo para proporcionar una visión general sobre la plataforma web ProRealSim con base en las experiencias y

Tabla 1 Temas, subtemas y códigos relacionados con la percepción de la plataforma ProRealSim

Análisis inductivo		
Tema	Subtemas	Códigos
Funcionamiento de la web	Funcionamiento del home	Adecuado
	Funcionamiento general	Sencillo Necesidad de mejora móvil
Contenido del medidor	Indicaciones específicas para el cálculo	Adecuadas/fácil lectura
	Indicadores	Pertinencia general Formación previa Subjetividad Pertinencia particular
	Fatiga del tasador	Número de ítems Número y ritmo de evaluaciones

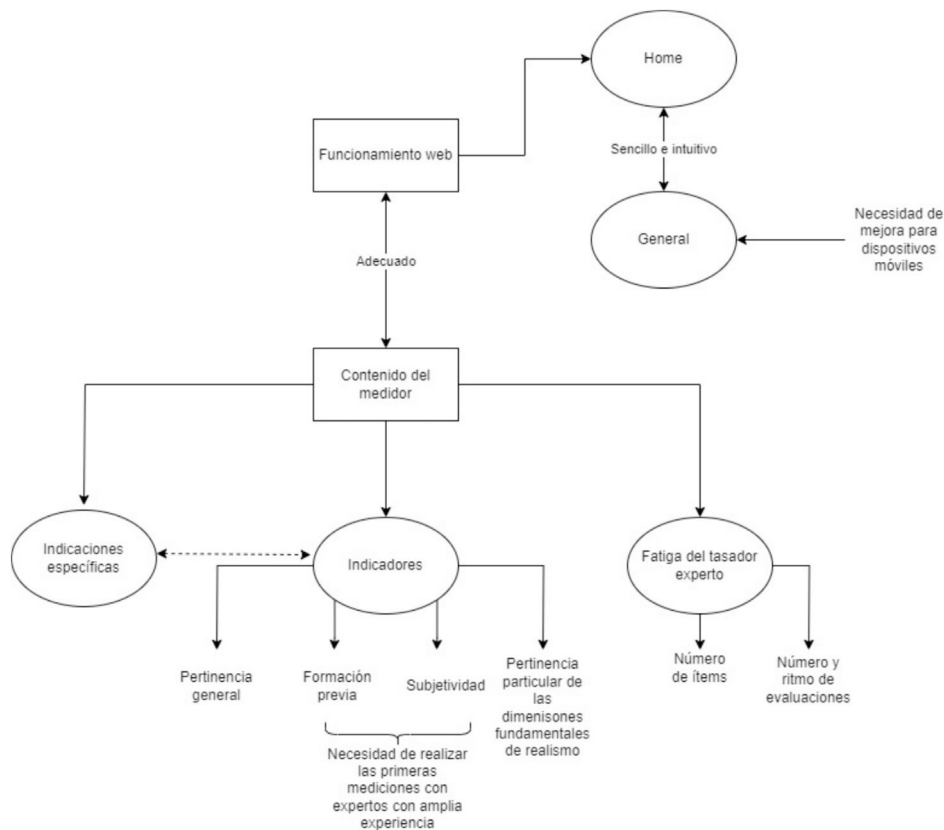


Figura 1 Establecimiento de relaciones entre temas (rectángulos), subtemas (círculos) y códigos (formato texto). Una vez analizados los datos, se establecen relaciones entre los subtemas y códigos basadas en las respuestas de los expertos, a partir de los temas de los que se partieron. Se observa que la página principal (home) es sencilla, pero requiere mejoras para su uso con dispositivos móviles. En cuanto al contenido del medidor, se encontró que las indicaciones específicas de la herramienta ProRealSim que deben ser evaluadas en cada simulación son pertinentes. No obstante, sería adecuado considerar indicaciones específicas dependiendo del escenario trabajado, para asegurar su relevancia. Por otro lado, para una correcta evaluación de los indicadores, es esencial que el evaluador tenga una formación previa y que, en sus primeras mediciones, cuente con la supervisión de expertos para evitar una posible subjetividad. Finalmente, es importante destacar la cuestión de la fatiga del evaluador. Se debe considerar el número de ítems y el ritmo de las evaluaciones, ya que un elevado número de ítems y evaluaciones en un corto período puede aumentar dicha fatiga.

perspectivas de los expertos que fueron entrevistados. En la [tabla 1](#) se muestran los distintos temas de los que se partió en la investigación y los subtemas y códigos que surgieron una vez se analizaron las entrevistas. Por su parte, en la [figura 1](#) se describen las relaciones entre los diferentes códigos que se establecieron una vez que se obtuvieron los resultados.

A continuación, en los siguientes subapartados se refieren los resultados de los 2 temas: funcionamiento web y contenido del medidor.

Funcionamiento web

Sobre el funcionamiento web surgieron 2 subtemas principales relacionados con el funcionamiento del home y el funcionamiento general de la web que alojó la herramienta ProRealSim.

En el funcionamiento del home de la página web las valoraciones fueron positivas. Según el experto M1 «(...) Realmente lo explicaba bien, facilitando la comprensión de lo que era el escenario, el simulador...».

En esta misma línea, para el funcionamiento general de la web, las percepciones de los expertos fueron similares, afirmando que era sencillo de utilizar: «Creo que era amigable el funcionamiento y que, cuando te acostumbras a utilizarla, es algo de fácil manejo y rápido» (F2).

Si bien es cierto, los expertos apuntaron la necesidad de realizar mejoras en la versión para dispositivos móviles:

«Yo unas cuantas mediciones las hice desde un dispositivo móvil y ahí reconozco que de amigable tenía cero (...) era bastante tedioso (...)» (F2).

Contenido del medidor

Respecto al tema del contenido del medidor ([tabla 1](#)), se tuvieron en cuenta 3 subtemas:

Subtema 1, *indicaciones específicas para el cálculo*: entendido como el análisis del conjunto de indicaciones que orientaron al usuario sobre el modo de utilizar la herramienta (indicación de cumplimentar todos los datos cualitativos para que se generen los cálculos, indicación de

introducir valores de 1 a 10, donde 1 era mínimo y 10 máximo, en todas las casillas que aplicasen, indicación de dejar en blanco la casilla del indicador que no aplicase e indicación de calificar con un cero el indicador que aplicase, pero no hubiera sido trabajado).

Subtema 2, *indicadores*: entendido como el análisis global de los ítems y, en particular, de aquellos que más polémica suscitaron para los evaluadores.

Subtema 3, *fatiga del tasador*: donde se debatió sobre las causas y modos de reducir el cansancio del evaluador.

Durante el análisis surgieron distintos códigos que se englobaron en los subtemas correspondientes. A continuación, se profundiza en cada uno de los subtemas generados.

Para el subtema *indicaciones específicas para el cálculo*, los expertos expresaron que eran útiles y claras. Para el subtema relacionado con los *indicadores*, se evaluaron las 3 dimensiones que contempló la herramienta (participante simulado, escenografía y simulador)⁹. En relación con la pertinencia general de los *indicadores*, asociados a las escalas de 1 a 10, los expertos expresaron que todos eran adecuados y fácilmente comprensibles.

«A mí, por lo menos, con la formación previa, no me dio problemas» (M1).

«(...) yo igualmente también había recibido la formación previa y creo que las etiquetas quedan bastante claras en lo referente a la precisión, la naturalidad y las 3 dimensiones fundamentales del realismo» (M2).

A pesar de que la percepción general de los indicadores fue adecuada, el grupo focal enfatizó en la necesidad de tener formación previa para poder realizar las mediciones de los indicadores de manera correcta en relación con cómo evaluarlos durante la simulación:

«Ando un poco perdido y sigo pensando que es porque tengo poca experiencia (...) tengo dificultad para poner el 0 y el 10» (M2).

«Coincido en que quizás el hacer muchas simulaciones te da ese punto más crítico» (F1).

Otro de los problemas mencionados por los expertos fue la subjetividad en la valoración. Indicaron que existe un gran componente *subjetivo* en el proceso de evaluación:

«No sé realmente en qué se basa un examinador en decir, bueno, esto es un 6 y estos son un 8» (M2).

En relación con el código de la pertinencia particular, el cual se centró en la relevancia de aspectos específicos para la evaluación de la simulación como los indicadores de tacto, olor, el nivel de detalle en el maquillaje y la peluquería, los expertos indicaron que algunas variables no eran relevantes en las simulaciones. Por ejemplo:

«Tanto olor, maquillaje, peluquería... me abrumaba al leer eso y era como, pero ¿qué hago con eso? ¿Qué puntúo?» (M1).

«Como dice M1, efectivamente hay ciertas cosas que, a pesar incluso de tener una persona real o no real, el tacto es difícilmente valorable en estos aspectos, ¿no? Y la peluquería, a grandes rasgos, tiene poca influencia en todo esto porque que tenga el pelo recogido, no recogido que tenga el pelo, alisado, liso, influye poco en el realismo, porque hay tantas diferencias, tanta variabilidad entre todos los tipos de personas que no creo que eso influya» (F2).

Sin embargo, salió a relucir que los indicadores podían ser o no útiles dependiendo de la simulación.

«Pero la vestimenta puede ser importante para el paciente en los escenarios de psicología(...)» F1.

En conjunto, los expertos enfatizaron que la complejidad de determinar la pertinencia y utilidad de los diferentes indicadores en el contexto de las simulaciones variaba en función de los casos evaluados, los indicadores, así como con la experiencia que tuviesen evaluando.

Otro aspecto de estudio fue la fatiga del experto a la hora de realizar las evaluaciones. Un aspecto analizado en este subtema fue que, el gran *número de ítems* generaba cierto agotamiento, principalmente cuando eran muy similares entre sí:

«de la emoción, la transición de las emociones, la improvisación con el alumno era como muy parecidos, y a mí eso me agotaba» (M1).

No solo el número de ítems, si no el número y ritmo de las evaluaciones fue un aspecto importante a la hora de aumentar la *fatiga del tasador*:

«A ver, no voy a mentir, fue cansado, pero es verdad que yo creo que también influye que fuera cansado por la cantidad de mediciones seguidas que hicimos» (F2).

Discusión

Los trabajos que han desarrollado herramientas para medir el realismo en entornos simulados se centran en consideraciones cuantitativas, como la validez estadística y en la aplicación de cuestionarios tradicionales. Wilson et al. y Sedlack et al.^{1-3,22} desarrollaron estudios en esta línea para entornos físicos y virtuales. Pese a la significación estadística de sus estudios, se echan de menos trabajos cualitativos que tomen en cuenta la percepción de los evaluadores con las herramientas aplicadas, sobre todo si se vinculan a evaluaciones online. En nuestro estudio, se pretendió ofrecer un análisis cualitativo de la experiencia de usuarios expertos en el uso de una herramienta online ya validada⁹.

El grupo focal emitió valoraciones positivas sobre el funcionamiento del *home* y general de la web. Estos criterios coinciden con los de Woit et al.²³ al plantear que las evaluaciones online resultan más precisas y realistas que las evaluaciones tradicionales. En concordancia con las ventajas de sencillez y acceso, y la disminución de costos que plantea el estudio de Orosa et al.²⁴, nuestros evaluadores consideraron que ProRealSim permitió la mejor exploración de los indicadores, resultando fácilmente comprensibles por los expertos, mejorando la retroalimentación y las oportunidades prácticas de aplicación^{23,25}. El estudio de Orosa et al.²⁴ también plantea una desventaja importante, pues, aunque existen herramientas de evaluación en Internet, es necesario que los evaluadores tengan un conocimiento profundo de su dinámica y metodología de uso para poder utilizarlas adecuadamente, lo que apoya nuestros resultados. También Williams et al.²⁶ consideraron que las evaluaciones online entrañan un desafío porque necesitan un desarrollo previo y, por lo tanto, deben ser más trabajadas y revisadas. Con esta misma percepción, el grupo focal de nuestro estudio consideró la necesidad de formación previa y de realizar las primeras evaluaciones acompañados

por docentes con experiencia. Coincidiendo con las conclusiones del estudio de Giebel et al.²⁷, nuestros evaluadores indicaron que la experiencia es factor determinante para aplicar los indicadores de manera efectiva y segura. Además, la percepción y la interpretación personal de cada evaluador influyeron en los puntajes asignados. A pesar de la experiencia formativa y del uso previo de las herramientas, esta subjetividad a veces resultó en diferencias en las calificaciones, incluso para un mismo caso de simulación evaluado. Este argumento coincide con lo considerado por Cate²⁸ al cuestionar la viabilidad de lograr una evaluación totalmente objetiva, argumentando que los juicios subjetivos de los expertos deben ser reconocidos y aceptados como parte integral de la evaluación.

Con base en los hallazgos, nuestro estudio aportó la recomendación de contar con, al menos, 2 tasadores que, a partir de objetivos de aprendizaje consensuados, evaluaran cada escenario. Malet²⁹ sugiere, igualmente, que la multiplicidad de criterios mejora el rigor y la exhaustividad de la evaluación. En este sentido, nuestro grupo focal propuso que el resultado final de estas evaluaciones fuese un índice de realismo obtenido a partir del promedio de las evaluaciones de cada uno de los tasadores involucrados, lo que establece un estándar numérico que comprende el criterio de todos.

Por otro lado, se analizó la pertinencia de ciertos indicadores (sonido, olor, maquillaje) que no tuvieron tanta presencia en los escenarios evaluados, pero el consenso general fue no eliminar ninguno, porque podrían ser importantes, según el contexto. Estos postulados estuvieron de acuerdo con los criterios de Chung³⁰ y Najj³¹.

Dado que no se llegó a un consenso para eliminar ningún ítem, se reconoció la importancia de evitar «maratones evaluativos» con el fin de prevenir el agotamiento del tasador. Además, se consideró la ventaja de realizar mediciones previas de la precisión, en los casos en que fuese posible, evitando realizar muchas evaluaciones en una sola jornada para mitigar el sesgo por olvido y el cansancio del tasador. Esta estrategia es una adaptación del postulado de Palacios³², que propone trabajar con versiones reducidas del cuestionario, por lo que en nuestro estudio se sugirió adelantar la evaluación previa de los indicadores referidos a la verosimilitud física o estructural (precisión) que se pueden evaluar antes del comienzo de la escena.

Otro elemento de consenso del equipo fue el considerar el realismo evaluado por el profesor como una variable vinculada pero no dependiente del desempeño del alumno. Este supuesto es contrario a literatura revisada, que otorga la mayor atención a la percepción del realismo del alumno, como se referencia en varios estudios^{33–35}, por lo que se recomendó estimular el uso de herramientas de este tipo, como ProRealSim o similares.

Conclusión

El estudio realizado permitió consolidar y recomendar procesos y métodos a la hora de evaluar cuantitativamente el realismo. Entre sus principales aportes podemos mencionar que ofrecer capacitación previa a los tasadores favorecería el uso riguroso de la herramienta, aportando a la eficacia y objetividad. Adicionalmente, apoyarse en múltiples evaluadores contribuiría para la reducción de la

subjetividad, favoreciendo perspectivas grupales en detrimento de opiniones personales y permitiendo que evaluadores sin mucha experiencia se apoyaran en evaluadores experimentados, consolidando una cultura de la evaluación. Valorar opciones de simplificar el cuestionario dependiendo del contexto a que se aplique podría reducir la fatiga del evaluador en casos de muestras grandes, favoreciendo muestras amplias y también la consistencia entre mediciones. Finalmente, realizar más estudios evaluativos de este tipo beneficiaría la cultura crítica en simulación clínica, evitando que el realismo continúe siendo una variable inaprensible y pobremente estudiada.

Financiación

Este trabajo ha sido totalmente financiado por la Beca de Investigación de la Sociedad Española de Simulación Clínica y Seguridad del Paciente (SESSEP), concedida en el año 2023, España. ID: UPD6278XWC26ECN8.

Conflicto de intereses

Los autores declaran que no tienen conflictos de intereses.

Ética de Publicación

Declaramos que se han tenido en cuenta las instrucciones para los autores. Además, confirmamos que nuestro estudio ha seguido los más altos estándares éticos en la investigación, obteniendo el consentimiento informado de todos los participantes y respetando su privacidad.

Bibliografía

1. Wilson E, Hewett DG, Jolly BC, Janssens S, Beckmann MM. Is that realistic? The development of a realism assessment questionnaire and its application in appraising three simulators for a gynecology procedure. *Adv Simul.* 2018;3(1):1–7.
2. Sedlack RE, Baron TH, Downing SM, Schwartz AJ. Validation of a colonoscopy simulation model for skills assessment. *Off J Am Coll Gastroenterol ACG.* 2007;102(1):64–74.
3. Sedlack RE. Validation of computer simulation training for esophagogastroduodenoscopy: pilot study. *J Gastroenterol Hepatol.* 2007;22(8):1214–9.
4. Steinberg AD, Bashook PG, Drummond J, Ashrafi S, Zefran M. Assessment of faculty perception of content validity of PerioSim®, a haptic-3D virtual reality dental training simulator. *J Dent Educ.* 2007;71(12):1574–82.
5. Seagull FJ, Rooney DM. Filling a void: developing a standard subjective assessment tool for surgical simulation through focused review of current practices. *Surgery.* 2014;156(3):718–22.
6. Cook DA, Hatala R, Brydges R, Zendejas B, Szostek JH, Wang AT, et al. Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. *Jama.* 2011;306(9):978–88.
7. Paige JB, Morin KH. Simulation fidelity and cueing: a systematic review of the literature. *Clin Simul Nurs.* 2013;9(11):e481–9.
8. Tun JK, Alinier G, Tang J, Kneebone RL. Redefining simulation fidelity for healthcare education. *Simul Gaming.* 2015;46(2):159–74.
9. ProRealSim [Internet]. Madrid: Universidad Europea. [consultado 15 Jun 2023]. Disponible en: <https://prorealim.universidadeuropea.com/>

10. Coro-Montanet G, Pardo Monedero MJ, Sánchez Ituarte J, Wagner Porto Rocha H, Gomar Sancho C. Numerical assessment tool to measure realism in clinical simulation. *Int J Environ Res Public Health*. 2023;20(3):2247.
11. Coro-Montanet G, Oliva-Fernández Ó, Sánchez-Ituarte J, Pardo-Monedero MJ. Nuevo concepto y herramienta para evaluar objetivamente el realismo en simulación clínica. *Rev Latinoam Simul Clin*. 2023;5(1):30–7.
12. Hamstra SJ, Brydges R, Hatala R, Zendejas B, Cook DA. Reconsidering fidelity in simulation-based training. *Acad Med*. 2014;89(3):387–92.
13. Creswell JW. Research design: qualitative, quantitative, and mixed methods approaches. 3rd ed. [Internet]. Third Edit. *J Chem Inf Model*. 2009;53:1689–99.
14. Cleland JA. The qualitative orientation in medical education research. *Korean J Med Educ*. 2017;29(2):61–71.
15. Kivunja C, Kuyini AB. Understanding and applying research paradigms in educational contexts. *Int J High Educ*. 2017;6(5):26–41.
16. Tavakol M, Sandars J. Quantitative and qualitative methods in medical education research: AMEE Guide No 90: Part II. *Med Teach*. 2014;36(10):838–48.
17. Creswell JW. Research design: qualitative, quantitative, and mixed methods approaches. 3rd ed. Thousand Oaks, CA, US: Sage Publications, Inc; 2009. p. xxix, 260–xxix, 260.
18. Pope C, Mays N. Qualitative methods in health research. In: Pope C, Mays N, editors. *Qualitative research in health care*. 3rd ed. Oxford: Blackwell Publishing Ltd; 2006. p. 1–11. Disponible en: <https://doi.org/10.1002/9780470750841.ch14>
19. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol*. 2006;3(2):77–101.
20. Creswell JW, Poth CN. Qualitative inquiry and research design: Choosing among five approaches. 4th ed. Thousand Oaks: Sage Publications; 2016.
21. Muñoz Justicia J, Sahagún Padilla M. Hacer análisis cualitativo con Atlas. ti 7. Man uso Creat Commons Attrib. 2017:4.
22. Ribbens W, Malliet S, Van Eck R, Larkin D. Perceived realism in shooting games: towards scale validation. *Comput Hum Behav*. 2016;64:308–18.
23. Woit D, Mason D. Effectiveness of online assessment. *Proceedings of the 34th SIGCSE Technical Symposium on Computer Science Education*; 2003. p. 137–41.
24. Eiroá Orosa FJ, Fernández Pinto I, Pérez Sales P. Cuestionarios psicológicos e investigación en Internet: una revisión de la literatura. *An Psicol*. 2008;24(1):150–7.
25. Nutan MTH, Damps EL. Online assessments in pharmaceutical calculations for enhancing feedback and practice opportunities. *Curr Pharm Teach Learn*. 2014;6(6):807–14.
26. Williams D, Hricko M, Howell SL. Online assessment, measurement and evaluation: Emerging practices. Hershey, PA: Estados Unidos de América: Information Science Publishing; 2006.
27. Giebel GD, Speckemeier C, Abels C, Plescher F, Borchers K, Wasem J, et al. Problems and barriers related to the use of digital health applications: scoping review. *J Med Internet Res*. 2023;25, e43808.
28. ten Cate O, Regehr G. The power of subjectivity in the assessment of medical trainees. *Acad Med [Internet]*. 2019;94(3).
29. Malet O, Giacomone B, Repetto AM. Modelo de evaluación de la idoneidad didáctica de un proceso de estudio masivo en el contexto de la pandemia de SARS-CoV-2. *Bolema-Math Educ Bul*. 2022;36:625–49.
30. Chung C, Navarrete L, Florido J, Salamanca A, Torne P, López R, et al. Fiabilidad de un examen clínico objetivo estructurado para predecir la futura competencia profesional de estudiantes de medicina. *FEM Rev Fund Educ Méd*. 2019;22(3):137–43.
31. Nanji KC, Baca K, Raemer DB. The effect of an olfactory and visual cue on realism and engagement in a health care simulation experience. *Simul Healthc*. 2013;8(3):143–7.
32. Palacios CDF. Versión reducida del Athlete Burnout Questionnaire (ABQ): propiedades psicométricas preliminares. *Rev Psicol Deporte*. 2015;24(1):177–83.
33. Saus E-R, Johnsen BH, Saus JE-R, Eid J. Perceived learning outcome: the relationship between experience, realism and situation awareness during simulator training. *Int Marit Health*. 2010;62(4):258–64.
34. Aarkrog V. 'The mannequin is more lifelike': The significance of fidelity for students' learning in simulation-based training in the social-and healthcare programmes. *Nord J Vocat Educ Train*. 2019;9(2):1–18.
35. Brandão CFS, Collares CF, Marin HF. Student perception on high-fidelity simulation during the medical clerkship. *Stud Health Technol Inform*. 2013;192:960.