



ORIGINAL

Adaptación, validación e invarianza factorial de una escala para evaluar la calidad de los cursos virtuales desarrollados para estudiantes peruanos de ciencias de la salud durante la pandemia por la COVID-19

Claudio Intimayta-Escalante^{a,h,*}, Rubí Plasencia-Dueñas^b, Janeth N. Nuñez-Lupaca^c, Kevin Flores-Lovon^d, Nancy Ramirez-Bernuy^a, Giancarlo Pares-Ballasco^e, Natalia Barrientos-Odías^f y Ronald Castillo-Blanco^g

^a Facultad de Medicina de San Fernando, Universidad Nacional Mayor de San Marcos, Lima, Perú

^b Sociedad Científica de Estudiantes de Medicina, Universidad Nacional Pedro Ruiz Gallo - SOCIEM UNPRG, Lambayeque, Perú

^c Facultad de Medicina, Universidad Nacional Jorge Basadre Grohmann, Tacna, Perú

^d Facultad de Medicina, Universidad Nacional de San Agustín de Arequipa, Arequipa, Perú

^e Sociedad Científica de Estudiantes de Odontología, Universidad Mayor de San Marcos, Lima, Perú

^f Sociedad Científica de Estudiantes de Nutrición – SOCIENUT, Universidad Mayor de San Marcos, Lima, Perú

^g Facultad de Medicina, Universidad Peruana de Ciencias Aplicadas, Lima, Perú

^h Epi-gnosis Solutions, Piura, Perú

Recibido el 6 de mayo de 2024; aceptado el 22 de octubre de 2024

Disponibile en Internet el 30 de diciembre de 2024

PALABRAS CLAVE

Psicometría;
Estudio de validación;
Educación a distancia;
Cursos vituales;
Aprendizaje en línea;
Educación médica;
Estudiantes del área de la salud;
COVID-19;
Perú

Resumen

Objetivo: adaptar, validar y evaluar la invarianza factorial de una escala para evaluar la calidad de los cursos virtuales (CCV) desarrollados para estudiantes de ciencias de la salud (ECS).

Métodos: estudio transversal desarrollado con encuestas virtuales a ECS para evaluar la CCV del año académico 2020, así como características sociodemográficas y educativas. La validez, basada en el contenido, se estimó con la V_{Aiken} . La validez basada en la estructura interna se evaluó con los índices de bondad de ajuste (IBA) en el análisis factorial confirmatorio (AFC), considerando adecuado el índice de ajuste comparativo (CFI) y el índice de Tucker-Lewis (TLI) mayor a 0,90; así como valores de la raíz residual estandarizada cuadrática media (SRMR) y el error cuadrático medio de aproximación (RMSEA) menores a 0,08. La invarianza factorial se consideró adecuada cuando la $\Delta\text{CFI} \leq 0,010$ y la $\Delta\text{RMSEA} \leq 0,015$ entre modelos en distintos niveles. La confiabilidad se evaluó con el índice alfa y omega.

Resultados: de los 629 ECS incluidos, el 62,96% eran mujeres, el 49,76% estudiaba medicina humana y 316 otras áreas de la salud. La V_{Aiken} fue de 0,86 en el instrumento. El AFC mostró IBA adecuados (CFI = 0,956; TLI = 0,950; RMSEA = 0,073 y SRMR = 0,044). Se estimó la ΔCFI

* Autor para correspondencia.

Correo electrónico: claudio.intimayta@unmsm.edu.pe (C. Intimayta-Escalante).

($\leq 0,002$) y la $\Delta RMSEA$ ($\leq 0,004$) entre los modelos para las características evaluadas (sexo, lugar de residencia, carrera de estudio, etapa de formación y tipo de gestión universitaria). Los índices de confiabilidad fueron mayores a 0,85 en los dominios.

Conclusión: la escala de 30 ítems y 6 dominios para evaluar la CCV desarrollados para ECS en Perú mostró adecuadas evidencias de validez y confiabilidad. Además, demostró una adecuada invarianza factorial en diferentes contextos sociodemográficos y educativos de los ECS.

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

KEYWORDS

Psychometrics;
Validation;
Distance education;
Virtual courses;
Online Learning;
Medical education;
Students of the health area;
COVID-19;
Peru

Adaptation, validation, and factorial invariance of a scale to assess the quality of virtual courses developed for Peruvian health sciences students during the COVID-19 pandemic

Abstract

Objective: To adapt, validate and evaluate the factorial invariance of a scale that assess the quality of virtual courses (QVC) among health sciences students (HSS).

Methods: Cross-sectional study developed with virtual surveys to HSS to assess the QVC of the 2020 academic year, as well as sociodemographic and educational characteristics. Content-based validity was estimated with V_{Aiken} . Validity based on internal structure was assessed with the Goodness of Fit Indices (GFI) in the Confirmatory Factor Analysis (CFA), considering adequate the comparative fit index (CFI) and the Tucker-Lewis index (TLI) greater than 0.90, as well as values of the standardized root mean square residual (SRMR) and root mean square error of approximation (RMSEA) less than 0.08. Factor invariance was considered adequate when the $\Delta CFI \leq 0.010$ and the $\Delta RMSEA \leq 0.015$ between models at different levels. Reliability was assessed with the alpha and omega index.

Results: Of the 629 HSS included, 62.96% were women, 49.76% studied human medicine and 316 other health areas. The V_{Aiken} was 0.86 on the instrument. The CFA showed adequate GFI (CFI = 0.956, TLI = 0.950, RMSEA = 0.073 and SRMR = 0.044). The ΔCFI (≤ 0.002) and $\Delta RMSEA$ (≤ 0.004) between models were estimated for the characteristics evaluated (sex, place of residence, study career, stage of training and type of university management). Reliability indices were greater than 0.85 in the domains.

Conclusion: The 30-item scale and six domains to assess QVC developed for HSS in Peru showed adequate evidence of validity and reliability. In addition, it showed adequate factorial invariance in different sociodemographic and educational contexts of HSS.

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

Introducción

Los cursos virtuales se integraron gradualmente en la educación universitaria, pero la pandemia por la COVID-19 aceleró la migración completa hacia entornos digitales^{1,2}. Aunque la recepción inicial fue positiva en los estudiantes, muchos cursos virtuales carecían de un adecuado diseño o directrices específicas para abordar las necesidades educativas³, lo que afectó a los estudiantes de ciencias de la salud por la interrupción de su formación en centros de salud, laboratorios y comunidades, con repercusiones en sus conocimientos y habilidades ante una formación exclusivamente virtual⁴.

En Perú, la pandemia evidenció la necesidad de mejorar la formación de los profesionales de la salud⁵. Para continuar con la formación de profesionales, las facultades de ciencias de la salud se adaptaron a entornos digitales, aunque solo el 27% de las universidades peruanas tenían autorización para ofrecer programas semipresenciales antes de la pandemia, lo que limitó la experiencia en el desarrollo de cursos virtuales^{6,7}. Además, las desigualdades en el acceso a Internet, la falta de

recursos tecnológicos y el desconocimiento de plataformas virtuales dificultaron la implementación⁸.

Aunque existen recomendaciones para el desarrollo de cursos virtuales⁹, muchas facultades de ciencias de la salud no han evaluado los cursos virtuales implementados¹⁰. Por ello, el objetivo de esta investigación fue adaptar culturalmente y abordar la validez e invarianza factorial de una escala que permite evaluar la calidad de los cursos virtuales desarrollados para estudiantes de ciencias de la salud en Perú durante la pandemia de COVID-19.

Materiales y métodos

Diseño de investigación

Estudio transversal enfocado en la adaptación y validación de una escala creada para evaluar la calidad de cursos virtuales en la Universidad Nacional de Educación a Distancia (UNED) en España¹¹, con el propósito de aplicarla a

estudiantes de ciencias de la salud en Perú, uno de los países más afectados por la COVID-19 en el mundo, que tiene a Lima como la capital del país, donde se aglomeran la mayoría de las facultades de ciencias de la salud¹². El estudio se dividió en 3 fases para abordar la adaptación del instrumento, validación de la escala y evaluación de la invarianza factorial.

Población de estudio

Se aplicó una encuesta virtual a estudiantes universitarios peruanos con 18 o más años en diferentes profesiones de la salud (enfermería, farmacia y bioquímica, medicina humana, nutrición, obstetricia, odontología, psicología y tecnología médica). La encuesta evaluó la percepción de la calidad de los cursos virtuales del año académico 2020, desarrollados íntegramente con entornos virtuales debido a la pandemia. Además, se excluyeron estudiantes de ciencias de la salud que desarrollaron estos cursos sin aulas virtuales, aquellos que no completaron toda la encuesta o no se matricularon en todos los ciclos y cursos correspondientes al año académico evaluado.

Adaptación del instrumento

El proceso de adaptación al contexto peruano se realizó mediante la metodología Delphi, con un comité de 5 investigadores: 2 con experiencia en validación, 2 en educación médica y uno en docencia en escuelas de medicina. Los ítems de la versión adaptada fueron revisados en 2 rondas para evaluar su validez de contenido en 6 dominios: calidad general, metodología didáctica, diseño y navegación en la plataforma virtual, recursos multimedia y materiales académicos. La validez basada en el contenido se estimó con la V de Aiken, considerando adecuado un valor superior a 0,70¹³.

Prueba piloto

Después de la evaluación del contenido, se obtuvo una versión adaptada del instrumento y se realizaron pruebas piloto con estudiantes de Medicina Humana para evaluar la claridad de las preguntas y su aplicabilidad. Además, se incluyeron preguntas sobre características específicas de los estudiantes. Con base en los comentarios recibidos, se revisaron los ítems que culminó en una versión de 32 ítems, agregando los ítems 1 y 9.

Recolección de datos

Los datos se recopilaron con una encuesta virtual (<https://is.gd/Calidad20>) utilizando el software Research Electronic Data Capture (o REDCap), desde julio hasta diciembre de 2021, a través de redes sociales (Facebook, WhatsApp, Twitter e Instagram). Así se utilizó la escala para evaluar la calidad de los cursos virtuales que media el grado de acuerdo del participante con los enunciados, utilizando un formato de respuesta en escala de Likert entre 1 (totalmente en desacuerdo), 2 (en desacuerdo), 3 (ni de acuerdo o en desacuerdo), 4 (de acuerdo) y 5 (totalmente de acuerdo).

Análisis estadístico

El análisis estadístico se realizó en R Studio versión 4.2.2 con los paquetes «tableone», «psych», «lavaan», «semTools» y «semPlot». De esta forma, en el análisis descriptivo, se utilizaron medidas de tendencia central (media) y dispersión (desviación estándar) para las variables numéricas, además de medidas de resumen (frecuencia y porcentaje) para las variables categóricas. Se evaluó si el número de encuestados era suficiente con la estimación del error estándar relativo (ESR), considerado adecuada la cantidad de encuestados cuando el ESR era menor a 20%¹⁴. La adecuación de la muestra para el análisis factorial se evaluó con índices de Kaiser-Meyer-Olkin mayores a 0,70¹⁵.

La validez basada en la estructura interna se evaluó con el análisis factorial confirmatorio (AFC) utilizando el estimador de mínimos cuadrados ponderados con la media y varianza ajustada (WLSM en inglés). Los modelos se evaluaron con índices de bondad de ajuste (IBA) como el chi-cuadrado (χ^2) con sus grados de libertad (GL), el índice de ajuste comparativo (CFI), el índice de Tucker-Lewis (TLI), la raíz residual estandarizada cuadrática media (SRMR) y el error cuadrático medio de aproximación (RMSEA). Se consideró un buen ajuste con CFI y TLI mayor a 0,90, además del RMSEA y SRMR menor a 0,08^{16,17}.

La invarianza factorial se evaluó siguiendo el enfoque de Svetina et al.¹⁸, utilizando la estimación de máxima verosimilitud para escalas ordinales. El análisis se estructuró en 3 niveles de modelos: el modelo base o configuracional (M1), que establece la estructura factorial inicial; el modelo métrico (M2), que evalúa la equivalencia de las cargas factoriales; y el modelo escalar (M3), que analiza la equivalencia de las cargas factoriales e interceptos. La invarianza factorial se consideró adecuada con ΔCFI menor a 0,010 y $\Delta RMSEA$ menor a 0,015¹⁹.

La evaluación de la confiabilidad de la escala basada en la consistencia interna se realizó con el índice alfa (α) y omega (Ω) para evaluar la confiabilidad compuesta, considerándolo adecuado los valores mayores a 0,70^{20,21}.

Aspectos éticos

La investigación recibió aprobación de un comité de ética en Perú antes de las pruebas piloto y la recolección de datos realizadas de manera virtual. Los datos recolectados fueron analizados sin información que permita identificar a los estudiantes de ciencias de la salud participantes de esta investigación.

Resultados

Primera fase: adaptación de la escala y validez basada en el contenido

En la primera fase de la investigación, tras la evaluación inicial por los 5 miembros del comité de investigadores de la primera versión adaptada de la escala, se identificó un valor promedio de V de Aiken de 0,85 (95%IC: 0,80 a 0,96). Sin embargo, después de la primera prueba piloto se modificó la escala a una segunda versión con 36 ítems, que se aplicó 4 semanas después en una segunda prueba piloto. Debido a los

comentarios y sugerencias, se redujo la versión a 32 ítems (fig. 1). Esta versión fue revisada nuevamente por el comité de investigadores para evaluar la validez basada en el contenido, identificando un valor promedio de V de Aiken de 0,86 (95%IC: 0,72 a 0,92).

Segunda fase: aplicación de la escala a estudiantes de Medicina

Se utilizó la tercera versión de la escala con 32 ítems y, de las 337 respuestas obtenidas, se incluyó solo a 297 estudiantes de Medicina Humana. Debido a que algunos no eran estudiantes

de Medicina Humana ($n = 6$), estudiaban en escuelas profesionales sin aulas virtuales ($n = 9$), no respondieron a todas las preguntas ($n = 25$) y no se matricularon en todos los ciclos o cursos correspondientes ($n = 40$).

De los 297 estudiantes de Medicina Humana, más de la mitad eran mujeres (53,20%), con una edad promedio de 22,50 años ($DE = 3,44$). En cuanto a la distribución por año académico, el 24,58% eran de tercero ($n = 73$), el 22,55% de cuarto ($n = 64$), el 19,53% de quinto ($n = 58$), el 15,82% de segundo ($n = 47$), el 15,49% de primero ($n = 46$) y el 3,03% de sexto ($n = 9$). La mayoría de los estudiantes de Medicina eran de universidades públicas (69,02%), frente a

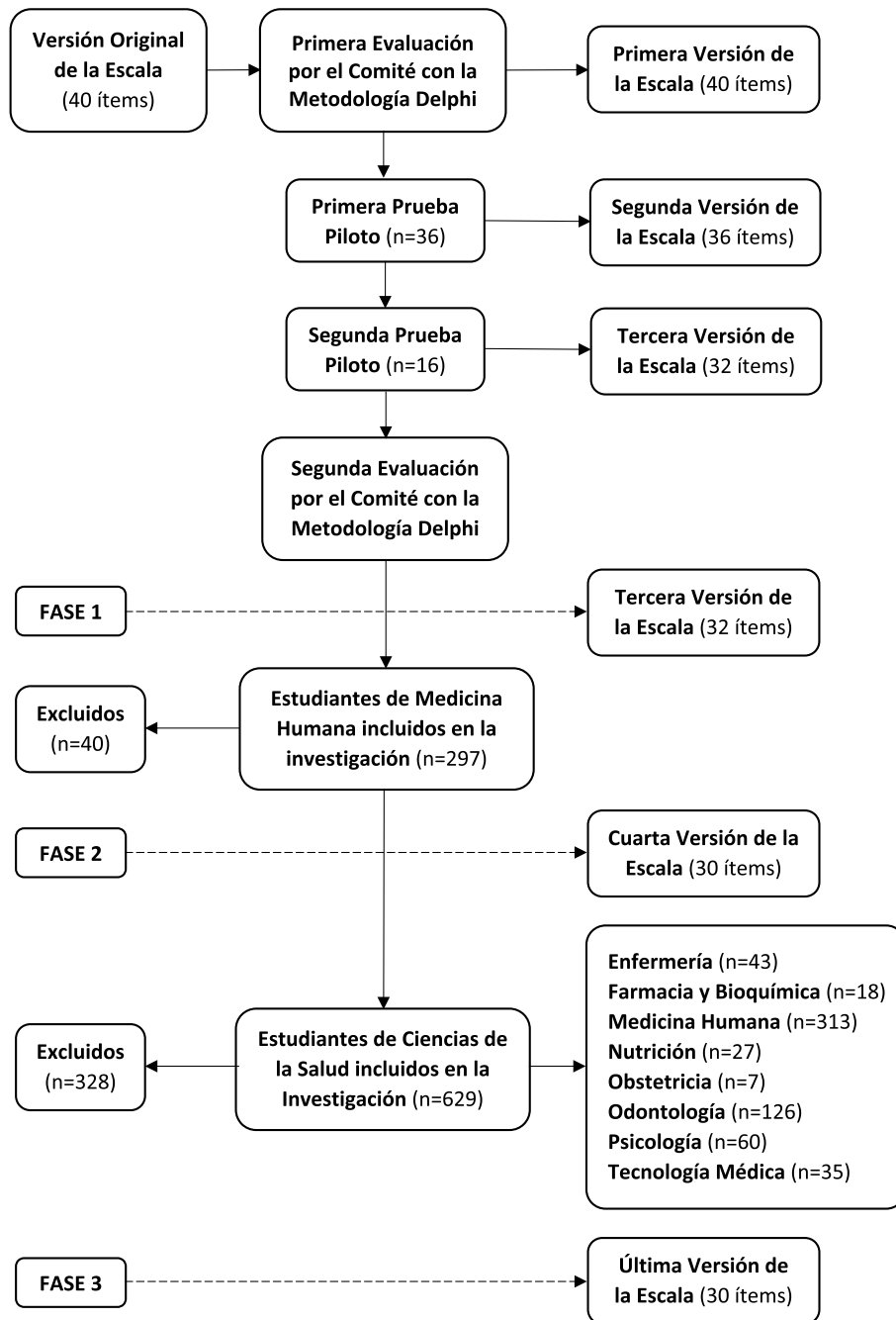


Figura 1 Flujograma del proceso de adaptación, validación e invarianza factorial de una escala para evaluar la calidad de los cursos virtuales desarrollados para estudiantes peruanos de ciencias de la salud durante la pandemia por la COVID-19.

Tabla 1 Distribución de la respuesta en los ítems de la escala que evalúa la calidad de los cursos virtuales desarrollados para estudiantes peruanos de ciencias de la salud durante la pandemia por la COVID-19

Nro.	Enunciado completo de cada ítem en la escala	Estudiantes peruanos de Medicina Humana (n = 297)					Estudiantes peruanos de ciencias de la salud (n = 629)				
		TD (%)	ED (%)	N (%)	DA (%)	TA (%)	TD (%)	ED (%)	N (%)	DA (%)	TA (%)
	<i>Los cursos virtuales...</i>										
1	...contaban con un programa o sílabo antes de comenzar	2,02	7,41	10,77	37,71	42,09	—	—	—	—	—
2	...tenían una calidad proporcional a la inversión	7,74	15,49	39,73	33,00	4,04	6,20	16,22	42,45	30,52	4,61
3	...tenían actividades variadas e interesantes	6,40	20,20	46,80	23,91	2,69	5,09	23,21	37,52	30,84	3,34
4	...promovieron la comprensión y el razonamiento de los temas	4,04	16,16	35,35	40,07	4,38	3,82	13,51	33,70	44,52	4,45
5	...eran organizados en el desarrollo de sus contenidos (desarrollando una introducción, los objetivos y los diferentes temas de cada capítulo)	4,71	12,79	30,98	43,77	7,74	4,13	11,76	28,30	47,22	8,59
6	...tenían contenidos exhaustivos y suficientes para el aprendizaje de los temas	3,37	17,17	39,73	35,69	4,04	3,66	14,94	36,09	40,38	4,93
7	...abordaron los contenidos de forma clara	3,03	11,78	39,06	39,73	6,40	2,70	10,97	36,09	45,47	4,77
	<i>Los docentes en los cursos virtuales...</i>										
8	...utilizaron una metodología de enseñanza que motivaba a reflexionar y/o preguntar sobre los temas	4,04	19,87	36,36	36,70	3,03	3,82	17,65	36,09	38,95	3,50
9	...encargaban una cantidad razonable de tareas	14,81	26,60	23,91	29,63	5,05	—	—	—	—	—
10	...desarrollaban evaluaciones adecuadas para el contenido desarrollado	4,71	15,82	26,60	48,48	4,38	4,45	14,47	26,87	49,28	4,93
11	...brindaban una adecuada retroalimentación para las tareas o evaluaciones desarrolladas	10,44	20,88	36,03	30,64	2,02	7,95	20,35	33,55	34,66	3,50
	<i>Los medios digitales usados para los cursos...</i>										
12	...facilitaron el proceso de enseñanza-aprendizaje (permitieron organizar grupos de trabajo, crear foros, brindar retroalimentación y otros)	2,69	11,78	33,33	45,45	6,73	2,86	11,13	27,34	50,72	7,95
13	...permitieron desarrollar ideas o interpretaciones novedosas que favorecieron la discusión sobre los temas	4,04	18,18	36,70	36,03	5,05	4,45	13,99	33,70	42,29	5,56
	<i>El aula virtual...</i>										
14	...tenía un formato agradable y equilibrado (adecuado tamaño de texto y calidad de imágenes).	4,04	12,46	27,27	47,81	8,42	4,61	10,49	25,76	49,76	9,38
15	...tenía una organización adecuada (jerarquización de tópicos e índice de contenidos)	3,03	12,12	24,58	51,85	8,42	3,18	11,29	22,26	54,37	8,90
16	...tenía encabezados, títulos y botones con buen diseño y funcionalidad.	2,36	12,79	24,24	50,17	10,44	2,86	10,33	21,78	54,21	10,81
17	...tenía páginas con una adecuada extensión (textos breves, hipervínculos, acceso rápido a menú principal o similar)	3,03	14,48	27,95	48,15	6,40	2,86	12,40	26,71	50,40	7,63
	<i>Las herramientas en el aula virtual...</i>										
18	...tenía páginas con una adecuada extensión (textos breves, hipervínculos, acceso rápido a menú principal o similar)	2,02	14,48	30,64	48,48	4,38	2,54	11,92	26,55	52,94	6,04
19	...permitieron una buena navegación entre los contenidos	2,36	12,12	33,33	45,12	7,07	2,38	10,33	28,14	51,51	7,63
20	...facilitaron la interacción entre usuarios (estudiantes y/o docentes)	2,69	20,20	32,32	40,74	4,04	3,34	16,85	29,57	44,20	6,04
21	...ayudaron a estudiar mejor los contenidos desarrollados en el curso	2,02	18,86	43,10	32,32	3,70	2,70	13,20	36,41	41,65	6,04
22	...ayudaron en la evaluación de los cursos (permitió desarrollar exámenes, visualizar las calificaciones o similar)	3,37	11,78	27,95	46,80	10,10	3,82	10,17	24,48	51,99	9,54
	<i>Los recursos multimedia (videos, simuladores y otros) usados en los cursos...</i>										

Tabla 1 (continuación)

Nro.	Enunciado completo de cada ítem en la escala	Estudiantes peruanos de Medicina Humana (n = 297)					Estudiantes peruanos de ciencias de la salud (n = 629)				
		TD (%)	ED (%)	N (%)	DA (%)	TA (%)	TD (%)	ED (%)	N (%)	DA (%)	TA (%)
23	...se caracterizaron por un diseño de calidad integrado a la estructura del aula virtual	6,06	20,54	34,34	35,35	3,70	4,29	17,49	33,70	40,22	4,29
24	...contaban con descripción general, transcripción y etiquetas del contenido	6,73	24,92	32,32	33,33	2,69	5,41	22,10	31,00	37,84	3,66
25	...eran variados e integrados a los temas desarrollados	6,06	19,53	34,01	37,71	2,69	4,29	15,10	31,96	43,72	4,93
26	...fueron de utilidad para las tareas asignadas <i>Las videoconferencias...</i>	6,06	17,51	31,31	41,08	4,04	4,77	13,04	29,25	47,06	5,88
27	...se programaban con suficiente anticipación	2,69	15,49	20,20	49,83	11,78	2,70	11,29	18,28	53,90	13,83
28	...se publicaban en el aula virtual para ser revisados posteriormente	7,74	20,20	21,21	39,06	11,78	5,56	15,10	21,46	45,15	12,72
29	...se realizaban a un ritmo adecuado <i>Los materiales académicos (libros, artículos o similares) usados en los cursos...</i>	5,72	18,18	30,64	38,05	7,41	4,29	13,20	27,50	46,26	8,74
30	...tenían contenidos comprensibles	2,69	7,74	25,25	59,93	4,38	2,38	6,36	22,73	62,48	6,04
31	...contenían información actualizada	2,02	11,78	27,27	49,83	9,09	2,07	9,70	24,64	53,58	10,02
32	...se adecuaban a los objetivos de aprendizaje	1,68	7,74	25,59	56,57	8,42	2,07	7,15	22,89	58,98	8,90

Nro.: número de ítem, TD: totalmente en desacuerdo, ED: en desacuerdo, N: ni de acuerdo o en desacuerdo, DA: de acuerdo, TA: totalmente de acuerdo.

92 estudiantes de universidades privadas. Por otro lado, solo el 27,95% vivía en Lima ($n = 93$).

Segunda fase: validez basada en la estructura interna y confiabilidad de la escala en estudiantes de Medicina

En el análisis estadístico descriptivo de los ítems de la escala aplicada a estudiantes de Medicina Humana, de un total de 160 puntos, el promedio del puntaje obtenido fue de 101,00 ($DE = 18,90$). En cuanto a la distribución de la respuesta en la escala de Likert, en cada ítem la mayoría expresó estar de acuerdo o totalmente de acuerdo (tabla 1).

En el AFC, se comparó los IBA de 3 posibles estructuras factoriales, según el artículo original, la versión con los 32 ítems y una versión con 30 ítems (tabla 2). De esta forma, se identificó un adecuado ajuste en el modelo de 30 ítems y 6 dominios ($\chi^2 = 657,66$; $GL = 390$; $CFI = 0,99$; $TLI = 0,99$; $RMSEA = 0,05$ y $SRMR = 0,05$). Asimismo, la correlación entre los dominios de este modelo de 30 ítems estuvo en el rango de 0,62 y 0,90, siendo este último valor, la correlación entre los dominios de calidad general y metodología didáctica. Mientras que en las cargas factoriales los valores estuvieron entre 0,55 y 0,93, con errores de varianza entre 0,14 y 0,68 (anexo 1).

En la evaluación de la confiabilidad del modelo de 30 ítems, se identificó un índice alfa y omega de 0,96, siendo que la confiabilidad de los 6 dominios de la escala tenía índices alfa mayores a 0,85 (0,87; 0,87; 0,90; 0,88; 0,91 y 0,86, respectivamente para cada dominio de la escala), y tenía índices omega mayores a 0,85 (0,88; 0,87; 0,90; 0,88; 0,91 y 0,86, respectivamente para cada dominio de la escala).

Además, se estimaron valores del ESR para los 32 ítems entre 1,30 y 2,03%, mientras que para el puntaje total fue 1,18%. El valor del índice de Kaiser-Meyer-Olkin o KMO global

fue de 0,95 y el rango de valores para los ítems fue similar (0,92 a 0,97).

Tercera fase: aplicación de la escala a estudiantes de ciencias de la salud

Luego de la validación de la escala con 30 ítems, se realizó la evaluación en estudiantes peruanos de ciencias de la salud, y de las 957 respuestas obtenidas hasta esta fase del estudio, se incluyó solo a 629 estudiantes, debido a que algunos no eran estudiantes de ciencias de la salud ($n = 18$), estudiaban en escuelas profesionales sin aulas virtuales ($n = 49$), no respondieron a todas las preguntas ($n = 99$) y no se matricularon en todos los ciclos o cursos correspondientes ($n = 162$).

De los 629 estudiantes de ciencias de la salud, el 62,96% eran mujeres con una edad promedio de 22,44 años ($DE = 3,52$). La mitad de los estudiantes encuestados eran de Medicina Humana (49,76%) (fig. 1). En cuanto a la etapa de formación, el 61,69% se encontraba en la etapa de formación avanzada (tercer año en adelante). El 79,33% eran de universidades públicas y el 55,33% eran de Lima (tabla 3).

Tercera fase: validez basada en la estructura interna y confiabilidad de la escala en estudiantes de ciencias de la salud

En el análisis estadístico descriptivo de los ítems de la escala aplicada a los estudiantes peruanos de ciencias de la salud, de un total de 150 puntos el promedio del puntaje obtenido fue de 104,00 ($DE = 19,60$). En cuanto a la distribución de la respuesta en la escala de Likert, en cada ítem la mayoría expresó estar de acuerdo o totalmente de acuerdo (tabla 1).

En el AFC del modelo con 30 ítems aplicado a los estudiantes de ciencias de la salud, se identificaron cargas

Tabla 2 Comparación de diferentes modelos de la escala que evalúa la calidad de los cursos virtuales durante la pandemia por la COVID-19

Modelo	Nro. de factores	Detalle de la correspondencia entre ítems y dominios propuestos	χ^2 (GL)	CFI	TLI	RMSEA	SRMR
<i>Basado en versión original de Santoveña S.^a</i>	3	<ul style="list-style-type: none"> Calidad del entorno (ítems 14, 15, 16, 17, 18, 19, 20, 21 y 22) Calidad de la metodología didáctica (ítems 2, 3, 4, 5, 6, 7, 8, 10, 11, 12 y 13) Calidad técnica (ítems 23, 24, 25, 26, 27, 28, 29, 30, 31 y 32) Calidad general y metodología didáctica (ítems 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 27, 28 y 29) 	1518.275 (402)*	0,909	0,901	0,100	0,071
<i>Basado en modelo de AFE con todos los ítems^b</i>	5	<ul style="list-style-type: none"> Diseño de la plataforma virtual (ítems 14, 15, 16 y 17) Recursos multimedia (ítems 23, 24, 25 y 26) Materiales académicos (ítems 30, 31 y 32) Navegación en la plataforma virtual (ítems 18, 19, 20, 21 y 22) Calidad general (ítems 2, 3, 4, 5, 6 y 7) Metodología didáctica (ítems 8, 10, 11, 12, 13, 27, 28 y 29) 	726.275 (395)*	0,954	0,949	0,072	0,051
<i>Estructura actual para AFC</i>	6	<ul style="list-style-type: none"> Diseño de la plataforma virtual (ítems 14, 15, 16 y 17) Navegación en la plataforma virtual (ítems 18, 19, 20, 21 y 22) Recursos multimedia (ítems 23, 24, 25 y 26) Materiales académicos (ítems 30, 31 y 32) 	657.659 (390)*	0,997	0,996	0,048	0,048

AFC: Análisis Factorial Confirmatorio; AFE: Análisis Factorial Exploratorio; CFI: Índice de Ajuste Comparativo; GL: Grados de Libertad; Nro.: número; RMSEA: Error Cuadrático Medio de Aproximación; SRMR: Raíz Residual Cuadrática Media Estandarizada; TLI: Índice de Tucker-Lewis; χ^2 : Chi-cuadrado.

^a Modelo basado en la estructura original de Santoveña Casal SM¹¹.

^b Modelo basado en la estructura del análisis factorial exploratorio identificado en el pre-print de Intimayta-Escalante C, et al.²².

* Las estimaciones tienen un valor p menor a 0,01.

Tabla 3 Validez basada en la estructura interna de la escala que evalúa la calidad de los cursos virtuales desarrollados para estudiantes peruanos de ciencias de la salud durante la pandemia por la COVID-19 según características sociodemográficas y educativas

Modelos	Índices de bondad de ajuste				
	χ^2 (GL)*	CFI	TLI	RMSEA	SRMR
<i>¿Qué carrera estudia?</i>					
Medicina humana (n = 313)	965.492 (390)	0,959	0,954	0,069	0,047
Otras carreras (n = 316)	1113.128 (390)	0,955	0,950	0,077	0,056
<i>¿Lugar de residencia?</i>					
Lima (n = 348)	1085.092 (390)	0,954	0,949	0,072	0,051
Otras regiones (n = 281)	952.822 (390)	0,961	0,957	0,072	0,049
<i>¿Etapas de formación?</i>					
Entre primer o segundo año (n = 388)	916.094 (390)	0,956	0,951	0,075	0,059
Desde tercer año (n = 241)	1179.169 (390)	0,956	0,951	0,072	0,047
<i>¿Sexo del encuestado?</i>					
Mujeres (n = 396)	1134.926 (390)	0,965	0,961	0,070	0,046
Hombres (n = 233)	841.385 (390)	0,952	0,946	0,071	0,057
<i>¿Tipo de gestión universitaria?</i>					
Gestión pública (n = 499)	1365.550 (390)	0,960	0,955	0,071	0,046
Gestión privada (n = 130)	692.079 (390)	0,948	0,942	0,077	0,067

CFI: Índice de Ajuste Comparativo; GL: Grados de Libertad; RMSEA: Error Cuadrático Medio de Aproximación; SRMR: Raíz Residual Cuadrática Media Estandarizada; TLI: Índice de Tucker-Lewis; χ^2 : Chi-cuadrado.

* El valor de p fue menor a 0,050 en todos los grados de libertad.

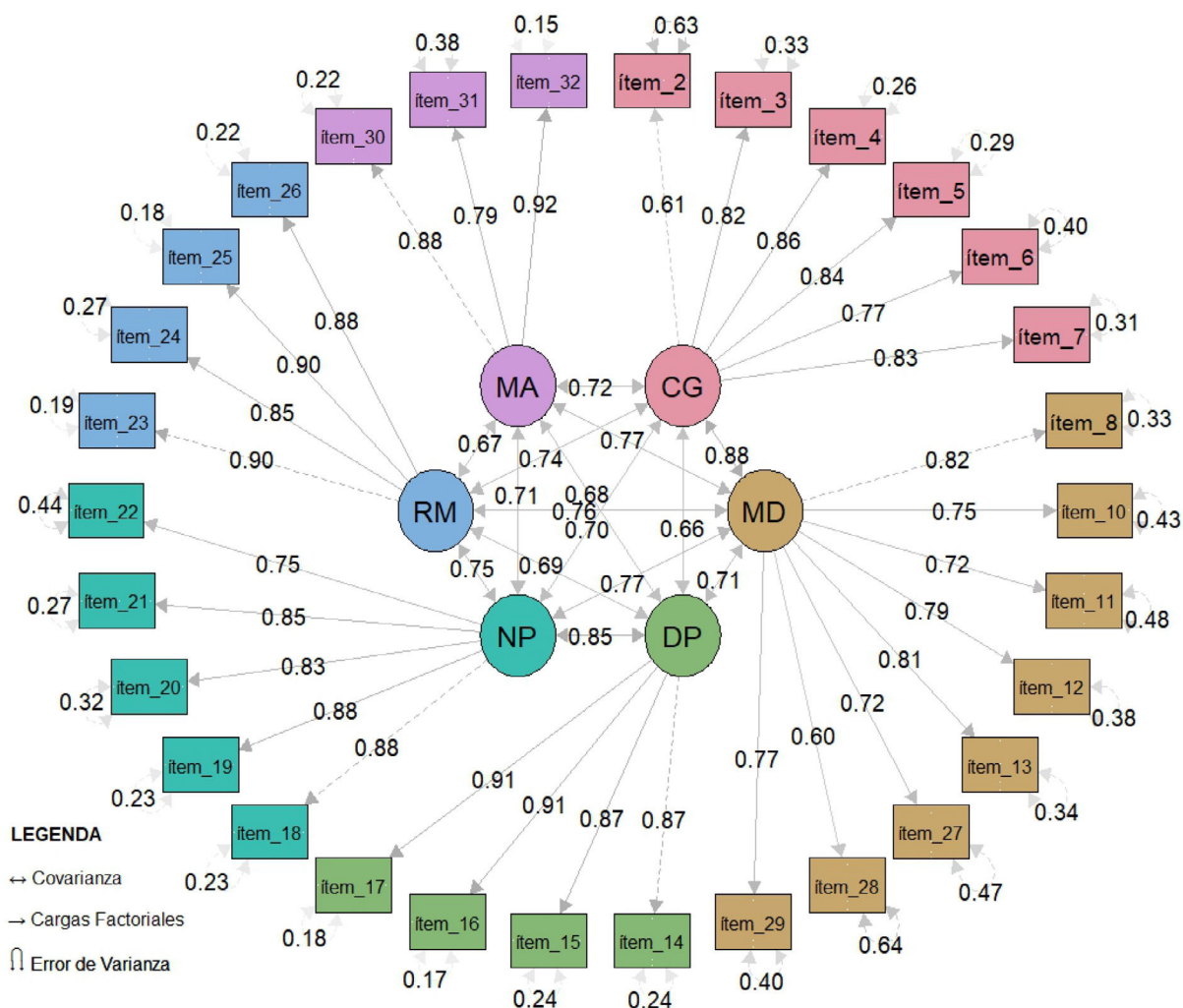


Figura 2 Modelo factorial de la escala para evaluar la calidad de los cursos virtuales desarrollados para los estudiantes peruanos de ciencias de la salud durante la pandemia por la COVID-19. **CG:** Calidad General; **DP:** Diseño de la Plataforma Virtual; **MA:** Materiales Académicos; **MD:** Metodología Didáctica; **NP:** Navegación en la Plataforma Virtual; **RM:** Recursos Multimedia.

factoriales entre 0,60 y 0,92 y covarianzas entre 0,66 y 0,88 (fig. 2), teniendo adecuados IBA ($\chi^2 = 1690,33$; $DF = 390$; $CFI = 0,956$; $TLI = 0,950$; $RMSEA = 0,073$ y $SRMR = 0,044$). Además, al estratificar a los participantes según las características evaluadas para la invarianza (sexo, lugar de residencia, carrera de estudio, etapa de formación y tipo de gestión universitaria), se identificó un adecuado ajuste para cada escenario evaluado (tabla 3).

En la evaluación de la confiabilidad se identificó un valor del índice alfa y omega de 0,96; que se mantuvo igual al evaluar la confiabilidad según la carrera estudiada, el lugar de residencia, la etapa de formación y el tipo de gestión universitaria. Sin embargo, los índices fueron mayores en estudiantes mujeres ($\alpha = 0,97$ y $\Omega = 0,97$), en comparación con estudiantes hombres ($\alpha = 0,95$ y $\Omega = 0,96$).

Tercera fase: invarianza factorial de la escala en estudiantes de ciencias de la salud

En la evaluación de la invarianza factorial, se identificaron adecuados valores de $\Delta CFI (\leq 0,002)$ y $\Delta RMSEA (\leq 0,004)$, que

disminuyeron entre los modelos restrictivos probados para todas las características evaluadas. Aunque en características como el lugar de residencia y la etapa de formación se encontraron valores más altos de ΔCFI y $\Delta RMSEA$, en comparación con características como la carrera de estudio, el tipo de gestión universitaria y el sexo del encuestado (tabla 4).

Discusión

Esta investigación adaptó y validó una escala para evaluar la calidad de los cursos virtuales dirigidos a estudiantes de ciencias de la salud en Perú durante el primer año de la pandemia de COVID-19. La versión adaptada y validada en el contexto peruano presenta mayor validez y confiabilidad que otros instrumentos²³⁻²⁶. Esto se debe a que muchos estudios solo reportan la confiabilidad de las preguntas sin detallar el proceso de validación del contenido o el constructo evaluado²³⁻²⁷.

La adaptación mediante la metodología Delphi garantizó una adecuada contextualización. La validez basada en el

Tabla 4 Invarianza factorial de la escala que evalúa la percepción de la calidad de los cursos virtuales desarrollados para estudiantes peruanos de ciencias de la salud durante la pandemia por la COVID-19

Modelo	Índices de bondad de ajuste						
	χ^2 (GL) *	CFI	Δ CFI	TLI	RMSEA	Δ RMSEA	SRMR
<i>Invarianza factorial según la carrera que estudia</i>							
M1	2086.206 (780)	0,956		0,951	0,073		0,051
M2	2150.860 (840)	0,957	0,001	0,955	0,070	0,003	0,051
M3	2136.348 (864)	0,958	0,001	0,958	0,068	0,002	0,052
<i>Invarianza factorial según el lugar de residencia</i>							
M1	2039.623 (780)	0,957		0,952	0,072		0,050
M2	2062.524 (840)	0,959	0,002	0,957	0,068	0,004	0,050
M3	2053.728 (864)	0,960	0,001	0,959	0,066	0,002	0,050
<i>Invarianza factorial según la etapa de formación</i>							
M1	2064.238 (780)	0,957		0,952	0,072		0,052
M2	2070.098 (840)	0,959	0,002	0,958	0,068	0,004	0,052
M3	2062.069 (864)	0,960	0,001	0,960	0,067	0,001	0,052
<i>Invarianza factorial según el sexo del encuestado</i>							
M1	1945.018 (780)	0,957		0,952	0,072		0,052
M2	1956.475 (840)	0,959	0,002	0,958	0,068	0,004	0,052
M3	1929.069 (864)	0,960	0,001	0,960	0,067	0,001	0,052
<i>Invarianza factorial según el tipo de gestión universitaria</i>							
M1	1903.870 (780)	0,962		0,958	0,068		0,050
M2	1877.215 (840)	0,965	0,003	0,964	0,063	0,005	0,050
M3	1861.627 (864)	0,967	0,002	0,966	0,061	0,002	0,051

CFI: Índice Comparativo de Ajuste; GL: Grados de Libertad; M1: Modelo Configuracional; M2: Modelo Métrico; M3: Modelo Escalar; Δ CFI: Diferencia entre los Índices Comparativos de Ajuste; TLI: Índice de Tucker-Lewis; RMSEA: Error Cuadrático Medio de Aproximación; Δ RMSEA: Diferencia entre el Error Cuadrático Medio de Aproximación; SRMR: Raíz Residual Cuadrática Media Estandarizada; χ^2 : Chi-cuadrado.

* El valor p fue menor a 0,001 en todos los grados de libertad.

contenido aseguró la claridad, representatividad y relevancia de cada ítem¹³. El análisis factorial confirmó la correspondencia entre los ítems y los dominios propuestos. Los IBA demostraron que la estructura teórica de ítems y dominios se ajusta a la realidad. Así la escala para evaluar la calidad cuenta con más evidencias de validez y confiabilidad que otros instrumentos²⁸.

La estructura de dominios e ítems de esta escala difiere de modelos anteriores con mejores IBA que avalan la versión de esta investigación^{11,22}. No obstante, aunque los 2 primeros dominios de la escala están correlacionados, los índices de ajuste sugieren que no deben combinarse, debido a que el primer dominio evalúa la calidad general de los cursos virtuales^{9,10}, mientras que el segundo analiza la metodología didáctica de los cursos virtuales^{23,24}.

Esta investigación también demostró la invarianza factorial de la escala, lo que confirma su aplicabilidad en diferentes condiciones (profesión, lugar de residencia, etapa de formación, sexo y tipo de gestión universitaria). A diferencia de otros instrumentos que evalúan cursos virtuales en estudiantes de ciencias de la salud, esta investigación demostró que el constructo fue interpretado de manera homogénea por todos los encuestados, independientemente de las condiciones sociodemográficas y educativas evaluadas^{8,18}.

Evaluar la calidad de los cursos virtuales es un desafío por su naturaleza subjetiva⁹⁻¹¹, pero la escala en esta investigación ofrece una forma de evaluación más precisa. Esto permite abordar mejor la perspectiva de los estudiantes como usuarios del servicio educativo²⁹. De esta forma, con la recolección de estas perspectivas en confidencialidad, anonimato y con

instrumentos validados, la optimización de estos cursos permitirá determinar ámbitos específicos de mejora.

La escala considera el diseño y la navegabilidad de la plataforma virtual para los cursos, pero algunas facultades no emplearon aulas virtuales⁸. La evidencia de la interacción con estas plataformas y su impacto en el aprendizaje demuestra la importancia de las aulas virtuales²¹. Sin embargo, esto supuso una limitación para la aplicación de la escala, más cuando el uso de recursos multimedia y materiales académicos ya eran empleados desde antes de la pandemia por la COVID-19²⁶.

La escala descrita en esta investigación facilita la evaluación de la calidad de los cursos virtuales implementados progresivamente en facultades de ciencias de la salud^{8,26}. La posibilidad de implementarlo en cursos específicos con modificaciones ayudaría en evaluaciones específicas. Esto para reconocer áreas de mejora a través de indicadores específicos. Además, puede integrarse a cursos con actividades virtuales y presenciales, contribuyendo a una estrategia de mejora continua³⁰.

Esta investigación presentó limitaciones, principalmente por el sesgo de selección al difundir la encuesta por redes sociales, lo que restringió la muestra a estudiantes con acceso a Internet, sin representar adecuadamente a la población de las facultades de ciencias de la salud en Perú. También hubo sesgos de medición por la autoaplicación de la encuesta y un posible sesgo de recuerdo, ya que algunos participantes fueron encuestados mucho después del año académico 2020. Además, combinar ciertas profesiones en una sola categoría pudo afectar la coherencia de la invarianza entre grupos. Por ello, futuras investigaciones podrían evaluar la escala en una muestra más representativa

y explorar otros aspectos del desarrollo de los cursos virtuales no incluidos en el instrumento.

En conclusión, la escala de 30 ítems y 6 dominios para evaluar la calidad de los cursos virtuales dirigidos a estudiantes de ciencias de la salud en Perú evidenció adecuada validez y confiabilidad. Además, la invarianza factorial permite su aplicación en estudiantes con diversas características sociodemográficas (sexo y lugar de residencia) y educativas (carrera, etapa de formación y tipo de gestión universitaria), facilitando una evaluación precisa y promoviendo la mejora continua de dichos cursos.

Responsabilidades éticas

La investigación fue aprobada por un comité de ética en Perú. Así se encuestó a los participantes empezando con el consentimiento informado (<https://is.gd/Calidad20>).

Financiamiento

Los autores declaran que no han recibido ayudas específicas del sector público y privado para la realización de este trabajo.

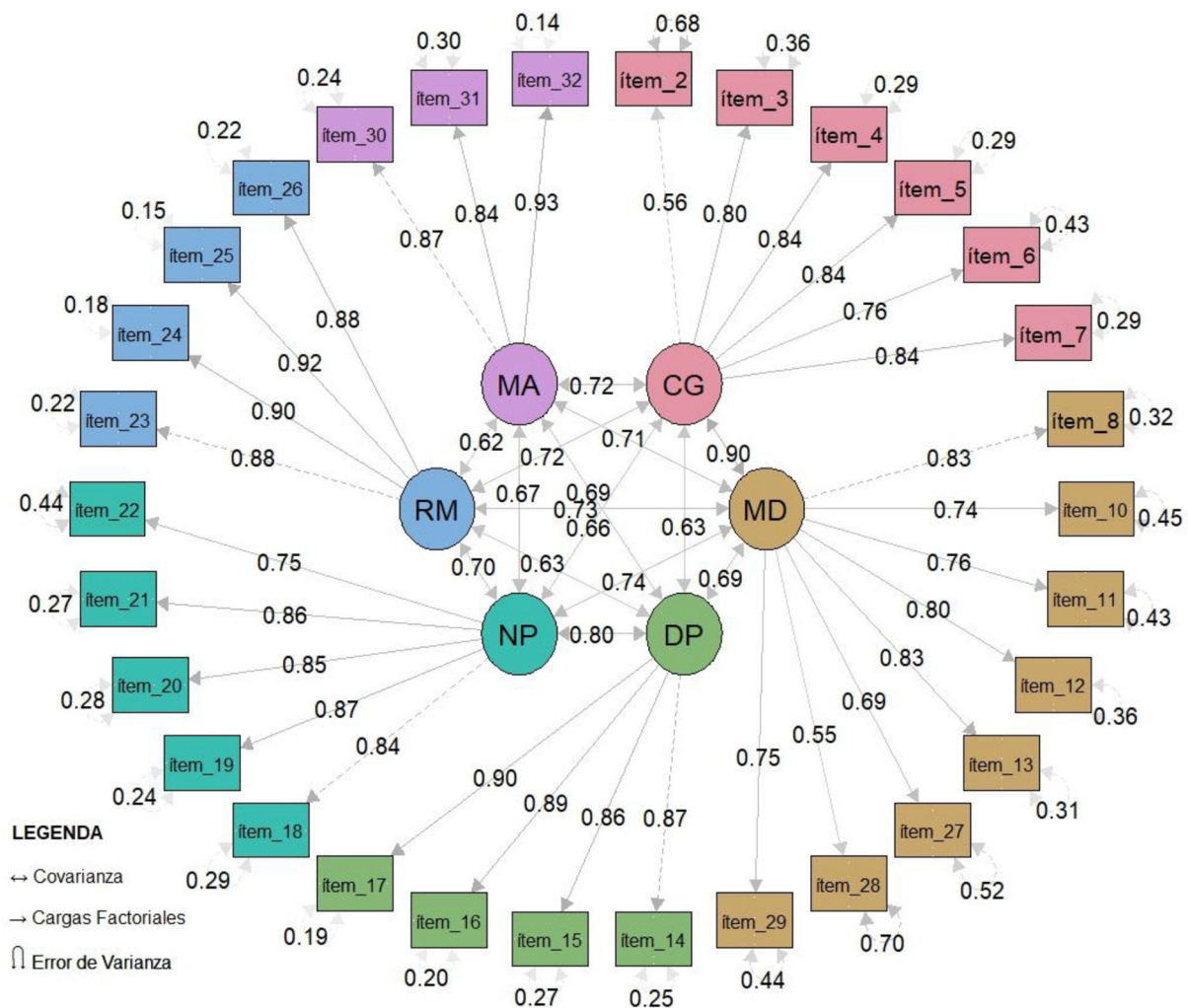
Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

Información adicional

Un informe preliminar de los resultados de esta investigación fue publicado como un preprint en la plataforma medrxiv: <https://www.medrxiv.org/content/10.1101/2022.01.08.22268928v1>

Appendix 1. Modelo factorial de la escala para evaluar la calidad de los cursos virtuales desarrollados para los estudiantes peruanos de medicina humana durante la pandemia por la COVID-19



CG: Calidad General, **MD:** Metodología Didáctica, **DP:** Diseño de la Plataforma Virtual, **NP:** Navegación en la Plataforma Virtual, **RM:** Recursos Multimedia, **MA:** Materiales Académicos.

Bibliografía

- Kim KJ, Kang Y, Kim G. The gap between medical faculty's perceptions and use of e-learning resources. *Med Educ Online*. 2017;22(1):1338504. <https://doi.org/10.1080/10872981.2017.1338504>.
- Rose S. Medical student education in the time of COVID-19. *JAMA*. 2020;323(21):2131–2. <https://doi.org/10.1001/jama.2020.5227>.
- Alsoufi A, Alsuyihili A, Msherghi A, Elhadi A, Atiyah H, Ashini A, et al. Impact of the COVID-19 pandemic on medical education: medical students' knowledge, attitudes, and practices regarding electronic learning. *PloS One*. 2020;15(11), e0242905. <https://doi.org/10.1371/journal.pone.0242905>.
- Beronja B., Bubnjevic T., Tasic R., Gasic M., Kulic L., Bogosavljevic I., et al., Comprehensive qualitative evaluation of the first ever full-time online course for medical students in one middle-income country, *Libyan J Med.*, **18** (1), 2023, 2258665, doi:10.1080/19932820.2023.2258665.
- Rodríguez M. La educación médica en tiempos del covid-19. *Rev Med Hered*. 2020;31:143–7. <https://doi.org/10.20453/rmh.v31i3.3802>.
- Sánchez-Duque J. Educación médica en tiempo de pandemia: el caso de la enfermedad por coronavirus 2019 (COVID-19). *Educ Méd*. 2020;21(4):259–60. <https://doi.org/10.1016/j.edumed.2020.05.005>.
- Ministerio de Educación del Perú, La universidad peruana: de la educación remota a la transformación digital. El sistema universitario frente al COVID-19 durante 2020 y 2021, [consultado 22 Mar 2024]. Disponible en: <https://bit.ly/3ZAUcqm>.
- Barrantes R., Burneo J. and Duffó D., No estábamos preparadas para las clases virtuales: la pandemia y la educación superior universitaria pública, 2022, Instituto de Estudios Peruanos, Documento de Trabajo – 285. [Consultado 22 Mar 2024], Disponible en: <https://bit.ly/3BnyUig>.
- Wasfy N., Abouzeid E., Nasser A., Ahmed S., Youssry I., Hegazy N., et al., A guide for evaluation of online learning in medical education: a qualitative reflective analysis, *BMC Med Educ.*, **21** (1), 2021, 339, doi:10.1186/s12909-021-02752-2.
- Mercado-Rey M., Cortez-Orellana S., Febres-Ramos R. Satisfacción estudiantil en una facultad de medicina por la virtualización de la enseñanza en el contexto de la pandemia de COVID-19. *FEM*. 2021;24(1):15–9. <https://doi.org/10.33588/fem.241.1106>.
- Santoveña S. Cuestionario de evaluación de la calidad de los cursos virtuales de la UNED. *Rev Educ Distancia*. 2010;10(25):1–22.
- Jimenez MM, Bui AL, Mantilla E, Miranda JJ. Human resources for health in Peru: recent trends (2007-2013) in the labour market for physicians, nurses and midwives. *Hum Resour Health*. 2017;15(1):69. <https://doi.org/10.1186/s12960-017-0243-y>.
- De Ventura-León J. regreso a la validez basada en el contenido. *Adicciones*. 2022;34(4):323–5. <https://doi.org/10.20882/adicciones.1213>.
- Barry D. Do not use averages with Likert scale data. Bookdown. 2017. [Consultado 22 Mar 2024]. Disponible en: <https://bookdown.org/Rmadillo/likert/>.
- Backhaus K, Erichson B, Gensler S, Weiber R, Weiber T. Factor analysis. *Multivariate Analysis*. Springer Gabler, Primera Edición; 2021. p. 381–450. https://doi.org/10.1007/978-3-658-32589-3_7.
- Bentler P. Comparative fit indices in structural models. *Psychol Bull*. 1990;107(2):238–46. <https://doi.org/10.1037/0033-2909.107.2.238>.
- Browne M, Cudeck R. Alternative ways of assessing model fit. *Sociol Methods Res*. 1992;21(2):230–58. <https://doi.org/10.1177/004912419201002005>.
- Svetina D, Rutkowski L, Rutkowski D. Multiple-group invariance with categorical outcomes using updated guidelines: an illustration using Mplus and the lavaan/semTools packages. *Struct Equ Model Multidiscip J*. 2020;27(1):111–30. <https://doi.org/10.1080/10705511.2019.1602776>.
- Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct Equ Model Multidiscip J*. 2002;9(2):233–55. https://doi.org/10.1207/S15328007SEM0902_5.
- Nunnally J, Bernstein I. *Psychometric Theory*. McGraw Hill. Tercera Edición; 1994;320–452. 007047849X, 9780070478497.
- Padilla MA, Divers J. A comparison of composite reliability estimators: coefficient Omega confidence intervals in the current literature. *Educ Psychol Meas*. 2016;76(3):436–53. <https://doi.org/10.1177/0013164415593776>.
- Intimayta-Escalante C, Plasencia-Dueñas R, Flores-Lovon K, Nuñez-Lupaca J, Chavez-Hermosilla M, Castillo-Blanco R. Adaptation and validation of a scale to evaluate the quality of virtual courses developed for medical students in Peru during the COVID-19 pandemic. *medRxiv*. 2022. <https://doi.org/10.1101/2022.01.08.22268928v1>.
- Goudarzi H, Onozawa M, Takahashi M. Impact of the Covid-19 pandemic and ensuing online teaching on pre-clinical medical education. *BMC Med Educ*. 2024;24(66). <https://doi.org/10.1186/s12909-023-04967-x>.
- Dost S, Hossain A, Shehab M, Abdelwahed A, Al-Nusair L. Perceptions of medical students towards online teaching during the COVID-19 pandemic: a national cross-sectional survey of 2721 UK medical students. *BMJ Open*. 2020;10(11), e042378. <https://doi.org/10.1136/bmjopen-2020-042378>.
- Elshami W, Taha M, Abuzaid M, Saravanan C, Al Kawas S, Abdalla ME. Satisfaction with online learning in the new normal: perspective of students and faculty at medical and health sciences colleges. *Med Educ Online*. 2021;26(1):1920090. <https://doi.org/10.1080/10872981.2021.1920090>.
- Stoehr F., Müller L., Brady A., Trilla A., Mähringer-Kunz A., Hahn F., et al., How COVID-19 kick-started online learning in medical education-The DigiMed study, *PloS One*, **16** (9), 2021, e0257394, doi:10.1371/journal.pone.0257394.
- Ventura-León J, Caycho-Rodríguez T, Talledo-Sánchez K. Satisfacción académica en estudiantes de Ciencias de la Salud antes y durante la pandemia COVID-19. *Rev Habanera Cienc Méd*. 2021;20(6):837606.
- Mohammadi S, Shohani M, Abedi L, Arizavi S, Rostami A. Quality of virtual education from the perspective of students and professors during the COVID-19 pandemic. *J Med Educ Dev*. 2022;15(46). <https://doi.org/10.52547/edcj.15.46.29>.
- Potdar P., Sharma J., Gupta U., Aggarwal D., Yadav Y., Quality and effectiveness of online medical education during COVID-19 outbreak: perception of medical students in Delhi NCR region, *J Datta Meghe Inst Med Sci Univ.*, **16** (2), 2021, 248–252, doi: 10.4103/jdmimsu.jdmimsu_77_21.
- Sahu PK, Dalcik H, Dalcik C, Gupta MM, Chattu VK, Umakanthan S. Best practices for effective implementation of online teaching and learning in medical and health professions education: during COVID-19 and beyond. *AIMS Public Health*. 2022;9(2):278–92. <https://doi.org/10.3934/publichealth.2022019>.