

Escala para documentos científicos de estandarización médica**Comprehensive content evaluation scale for scientific documents**

Mayelín de la Caridad Suárez González,^{1*} Silvia María Díaz Gómez,² Karell Rafael Vazquez Argote,³
Elisa Amparo Núñez Oduardo.⁴

¹Clínica Estomatológica 3 de Octubre, Las Tunas, Cuba, mcsuarez@ltu.sld.cu;

<https://orcid.org/0000-0002-1486-5481>

²Clínica Estomatológica La Vigía, Camagüey, Cuba, msilvia.cmw@infomed.sld.cu;

<https://orcid.org/0000-0002-6314-3434>

³Universidad de Ciencias Médicas Dr. Zoilo Marinello Vidaurreta, Las Tunas, Cuba; <https://orcid.org/0000-0003-3860-5702>

⁴Clínica Estomatológica 3 de Octubre, Las Tunas, Cuba, enoduardo@ltu.sld.cu;

<https://orcid.org/0000-0001-7171-8691>

*Autor para la correspondencia. Correo electrónico: mcsuarez@ltu.sld.cu

DOI: 10.70373/RB/2024.09.03.9

Resumen

El uso de escalas es de gran utilidad; se realizó una investigación de desarrollo y evaluación de marzo 2021 a mayo 2022 en la Clínica Estomatológica 3 de octubre de la provincia Las Tunas, Cuba, con el objetivo de crear una escala para la evaluación de documentos científicos de estandarización médica desde su contenido y realizar su validación por expertos. Se seleccionaron por muestreo simple aleatorio 15 profesionales de alto nivel científico, se consideró la metodología utilizada en documentos precedentes de la literatura científica y la validación fue realizada mediante el método Delphi. Se exploró la confiabilidad relacionada con la evaluación por diferentes personas, así como la validez de contenido con concordancia entre las evaluaciones de los jueces para el instrumento y la evaluación de la escala, mostrándose valores W de Kendall de 0.73 y 0.71, $p = 0,000$ respectivamente, aspecto que define a la escala creada como recomendada. Este instrumento permitirá evaluar documentos científicos desde su contenido, unificar criterios y elevar el grado de preparación de los profesionales de la salud en la contemporaneidad.

Palabras clave: escala de evaluación, escala de contenidos, validación, panel de expertos.

Abstract

The use of scales is very useful; a development and evaluation investigation was carried out from March 2021 to May 2022 at the 3 de Octubre dentistry clinic in the province of Las Tunas, Cuba; with the aim of creating a scale for the evaluation of scientific documents of medical standardization

from their content and validating them by experts. 15 professionals of high scientific level were selected by simple random sampling, the methodology used in previous documents from the scientific literature was considered and the validation was carried out using the Delphi method. The reliability related to the evaluation by different people was explored, as well as the content validity with agreement between the judges' evaluations for the instrument and the evaluation of the scale, showing Kendall's W values = 0.73 and 0.71; $p = 0,000$ respectively; and aspect that defines the scale created as recommended. This instrument will allow evaluating scientific documents from their content, unify criteria and raise the degree of preparation of health professionals in contemporary times.

Keywords: evaluation scale, content scale, validation, expert panel.

Introducción

La producción científica desde diversas disciplinas y escenarios ha crecido considerablemente en las últimas décadas. En este sentido, la evaluación de la calidad de los estudios se torna esencial por su transparencia, visibilidad, rigor e impacto de la producción y publicación científicas.¹

En lo que a ciencias sociales y humanas se refiere, se aprecia cada vez más la tendencia a ser lo más objetivo posible en la recolección de información de un determinado hecho; de esta manera, las escalas de medición han logrado un importante posicionamiento debido a que su utilización se viene generalizando no solo para evaluar las actitudes, sino también opiniones, comportamientos u otras características individuales.²

Las razones que impulsan al uso clínico de escalas, cuestionarios, tests, e índices pueden ser diversas, como son la ayuda al diagnóstico, el apoyo al umbral de tratamiento, la comunicación con otros profesionales, y el establecimiento de un referente poblacional.³

En la confección de las mismas, los expertos son necesarios, para comprender las verdaderas dimensiones del problema, desarrollar modelos alternativos, decidir qué datos utilizar o simplemente interpretar los resultados. En aquellas situaciones donde la solución de un problema o la adopción de decisiones sea muy importante, es mejor recoger opiniones de las personas con amplios conocimientos y experiencia contrastada en el área de interés y, por lo tanto, mejor cualificadas para responder a las cuestiones planteadas.⁴

De acuerdo con Castellano Serra,⁵ la necesidad de socializar resultados científicos de forma clara y transparente influye positivamente sobre la formación de nuevos conocimientos, y consecuentemente aumenta la confianza en las conclusiones cuando el estudio es realizado con rigor metodológico. Por lo tanto, la propagación de la utilización de las herramientas de evaluación metodológica y de

información, califica y legitima cada vez más la producción científica, principalmente en el campo de los procedimientos de la asistencia médica.

Esto evidencia la necesidad constante de elaborar y llevar a cabo procesos para la validación de instrumentos teóricos y/o prácticos que proporcionen una serie de oportunidades en la recolección de la evidencia científica.⁶ Razón por la que se propone crear una escala para la evaluación de documentos científicos de estandarización médica desde su contenido y realizar su validación por expertos.

Material y Métodos

Se realizó una investigación de desarrollo y evaluación de marzo 2021 a mayo 2022 en la Clínica Estomatológica 3 de Octubre de la provincia Las Tunas, Cuba. La misma consistió en la construcción y validación de un instrumento con su escala para evaluar documentos científicos de estandarización médica desde su contenido.

Universo y muestra del grupo evaluador

De un universo de 20 profesionales de alto nivel científico, se seleccionaron por muestreo simple aleatorio 15 como expertos: doctores en ciencias y especialistas de segundo grado de diferentes programas de especialización dentro de las Ciencias Estomatológicas, con categoría docente principal e investigativa, se consideró la experiencia de trabajo, integración a equipos interdisciplinarios, a los que se le solicitó el consentimiento informado. Se realizó una encuesta de autodeterminación de competencia de experto que permitió caracterizar a los mismos, obtener una autovaloración de su grado de conocimiento en el tema y la influencia que han tenido diversas fuentes de argumentación. Se determinó el coeficiente de competencia (K) para la selección de los expertos.⁴

Los valores finales de K obtenidos permitieron clasificar a los expertos en tres grupos:

- Alta influencia de todas las fuentes, que son los que obtuvieron un valor K superior a 0,8.
- Influencia media, los que alcanzaron un valor comprendido entre 0,7 y 0,8.
- Baja influencia, los que lograron una valoración inferior a 0,7.

Técnicas y procedimientos

Para la obtención de la información necesaria en la construcción del instrumento y escala integral de evaluación de contenido para documentos científicos de estandarización médica (ECDOCE), se consideró la metodología utilizada en otros documentos científicos precedentes en la literatura científica, que se resumen en cuatro pasos:^{2,7,8}

- Definición y caracterización del objeto de estudio: abordada en la introducción de la investigación.

• Selección de los indicadores: basados en la literatura científica, instrumentos teóricos desarrollados por el equipo multidisciplinario de la consulta del tercer nivel de atención en salud de Oclusión y Articulación Temporomandibular de Camagüey y la experiencia clínica e investigativa de profesionales miembros de la Cátedra multidisciplinaria para el estudio de la oclusión dentaria Plácido Ardanza Zulueta,⁸ así como de especialistas de experiencia en esta temática en el territorio. Se seleccionaron los ítems a tener en consideración.

• Construcción de los ítems: se tomó el tipo de pregunta cerrada, de selección única, con su escala nominal a la que se le otorgó una puntuación que permitió la realización de operaciones matemáticas, con las categorías: ausente (1), presente (2).

• Definición de la forma evaluativa del instrumento: se clasifica el resultado final del instrumento en una escala cualitativa ordinal, con el empleo de los cuartiles. Se consideraron las categorías: No recomendado, Parcialmente recomendado, Recomendado, Muy recomendado, en correspondencia con la puntuación obtenida.

Se diseñaron 16 ítems (variables del estudio), 15 correspondientes a los aspectos de observación del protocolo (variables cualitativas nominales) y una correspondiente a la forma de medición final del instrumento (escalas de cualitativa ordinal) que responden a las características del Sistema Nacional de Salud Pública en Cuba para la conformación del instrumento y su escala de evaluación para contenido. Estas fueron valoradas por el grupo de expertos para emitir un juicio valorativo mediante la escala Likert,⁹ de cinco categorías.

Las variables utilizadas fueron:

- Representatividad del problema de salud.
- Información teórico-práctica desde la evidencia científica.
- Conformación de un documento científico con la integración de saberes.
- Exposición de argumentos conciliadores pertinentes a la selección de los mejores procedimientos.
- Establecimiento de acciones clínicas con espacio para las individualidades.
- Inclusión de soluciones tangibles en el contexto histórico en que se implementan.
- Estructura y orden en las acciones o procedimientos médicos.
- Propuesta de alternativas específicas y cambio de procedimientos ante problemas identificados como limitaciones probables.
- Inclusión de elementos sustantivos propios de la educación.
- Prevención, diagnóstico y tratamiento del problema de salud identificado.
- Presentación y estratificación del método clínico.
- Empleo de tecnología y enfoque terapéutico desde el desarrollo científico-técnico.

- Proyección terapéutica en correspondencia con la etiología del fenómeno.
- Representación de técnicas y procedimientos clínicos con esquemas y/o algoritmos de trabajo.
- Distinción con precisión de las recomendaciones necesarias.
- Estipulación de los resultados relevantes de la discusión colectiva por etapas.
- Escala de evaluación del instrumento para evaluación de contenido.

Para la validación de la escala a partir de la experiencia de especialistas en el tema y complementada con las consideraciones individuales aportadas por los expertos, se aprobó en un primer momento el instrumento con cada uno de los ítems y luego se puso en consideración la evaluación de estos y la escala en cuestión. El análisis se realizó con el método Delphi.¹⁰ Se sometió a la valoración del panel de expertos mediante la escala Likert.^{7,9} La medición de los criterios se correspondió con cinco categorías que establecían una evaluación del uno al cinco: totalmente en desacuerdo con valor uno, en desacuerdo con valor dos, ni de acuerdo ni en desacuerdo con valor tres, de acuerdo con valor cuatro, muy de acuerdo con valor cinco.

Se formuló una interrogante abierta para que cada experto pudiese expresar su juicio teórico con respecto a posibles modificaciones, así como brindar sugerencias al respecto. Para las sesiones de trabajo se empleó la comunicación en línea y presencial.

El instrumento se consideró validado cuando los expertos emitieron una calificación por ítem que en la suma total correspondiera a la aprobación de la escala, categoría Recomendado correspondiente a una puntuación de mayor de 60 puntos.

Posteriormente para validar este instrumento se exploró:

Confiabilidad: es capacidad que tiene el instrumento de arrojar siempre los mismos resultados cada vez que se aplica a la misma unidad de observación.¹¹

Validez: es capacidad del test de arrojar los mismos resultados ante nuevas aplicaciones, distinguen los siguientes componentes:

- **Validez de apariencia:** se examinó a través de consulta a expertos si los ítems que componen el instrumento fueron correctamente elaborados.
- **Validez de contenido:** los expertos evaluaron el cumplimiento de los criterios de Moriyama¹² para cada uno de los ítems mediante la escala Likert.⁹
- **Validez de constructo:** se evaluó mediante la construcción de una hipótesis que contrasta el grado de asociación entre el valor obtenido por el instrumento creado y contenido del protocolo. Para ello se consideró la sumatoria de las puntuaciones obtenidas con la escala, resultando No recomendado si resultó menor de 19 puntos, o Muy recomendado, entre 27 y 30.

- **Validez de criterio:** el instrumento construido surgió dada la necesidad de interpretar su aplicabilidad a documentos científicos de estandarización médica.

Tanto para la construcción como para la validación del instrumento, los datos se obtuvieron de fuentes primarias con los que se confeccionó una base de datos. El procesamiento estadístico se realizó con medidas de resumen (mediana, desviación estándar), para la concordancia inter-expertos se emplearon, el test w de Kendall⁴ y el índice de concordancia V de Aiken.¹³ La correlación de elementos corregida se empleó para examinar la homogeneidad del aporte de los ítems y, para la medición de la confiabilidad el alfa de Cronbach,¹⁴ una confiabilidad del test en el 95 %, con el empleo del paquete estadístico SPSS 25.0 para Windows.

El cálculo del índice de concordancia V de Aiken¹³ (ecuación 1) con sus intervalos de confianza, para medir la concordancia inter-expertos, se realizó manualmente en la hoja de cálculo de Excel, (ecuaciones 2 y 3).

$$V_{Aiken} = \frac{\text{Mediana}-\text{Valor mínimo}}{\text{Valor Máximo}-\text{Valor mínimo}} \quad (1)$$

$$LIV_{Aiken} = \frac{2nkV+Z^2-Z\sqrt{4nkV(1-V)+Z^2}}{2(nk+Z^2)} \quad (2)$$

$$LSV_{Aiken} = \frac{2nkV+Z^2+Z\sqrt{4nkV(1-V)+Z^2}}{2(nk+Z^2)} \quad (3)$$

Para examinar la relevancia del coeficiente de correlación W de Kendall¹² se ejecutó el test chi-cuadrado¹⁵ (ecuación 4).

$$\chi^2 W = \frac{S}{\frac{1}{12}E^2(C^3-C)} \quad (4)$$

La calificación total de la escala presentó un margen de valores entre 15 y 75 puntos; para elaborar el criterio de aprobación se determinó el rango intercuartílico con los cuartiles 1 y 3 resultando:

- Menor de 30 puntos: No Recomendado
- 30-60 puntos: Recomendado con modificaciones
- Mayor de 60 puntos: Recomendado

Resultados

Es necesario incorporar instrumentos científicos que faciliten y normen la toma de decisiones interdisciplinarias a favor del desarrollo de la ciencia y propicien un método activo de aprendizaje desde la práctica. Con esta investigación se logró la creación de una escala para la evaluación de documentos científicos de estandarización médica desde su contenido.(figura 1)

ITEMS	Evaluación*
1 Representatividad del problema de salud.(según objetivo propuesto)	
2 Información teórico-práctica desde la evidencia científica	
3 Conformación de un documento científico desde la integración de saberes	
4 Exposición de argumentos conciliadores pertinentes a la selección de los mejores procederes	
5 Establecimiento de acciones clínicas con espacio para las individualidades	
6 Inclusión de soluciones tangibles en el contexto histórico en que se implementan	
7 Estructura y orden en las acciones o procederes médicos	
8 Propuestas de alternativas específicas y cambios de procederes ante problemas identificados como limitaciones probables	
9 Inclusión de elementos sustantivos propios de la educación, prevención, diagnóstico y tratamiento del problema de salud identificado.	
10 Presentación y estratificación del método clínico	
11 Empleo de tecnología y enfoque terapéutico desde el desarrollo científico- técnico	
12 Proyección terapéutica en correspondencia con la etiología del fenómeno	
13 Representación de técnicas y procedimientos clínicos con esquemas y/o algoritmos de trabajo	
14 Distinción con precisión de las recomendaciones necesarias	
15 Estipulación de los resultados relevantes de la discusión colectiva por etapas	
Total de puntos	
*: 1= Ausente; 2= Presente	
	Evaluación cualitativa
	Puntaje
	No recomendado
	< 19
	Escala: Parcialmente
	19-22
	Recomendado
	23-26
	Muy recomendado
	27-30

Figura 1. Escala integral de evaluación de contenido para documentos científicos de estandarización en las Ciencias Médicas. (ECDOCE)

En la tabla 1 se muestra la evaluación del coeficiente de competencia experta. En el presente estudio participaron 15 profesionales de la salud expertos con una puntuación superior a 0,8.

Tabla 1. Evaluación del coeficiente de competencia experta.

Expertos	Provincia	Grado científico u académico	Categoría docente principal	Kc ⁺	Ka ⁺⁺	K ⁺⁺⁺	Nivel
E- 1	Camagüey	Dr.C.	Titular	1	1	1	A
E- 2	Camagüey	MSc	Titular	0,9	0,85	0,85	A
E- 3	Camagüey	MSc	Auxiliar	1	0,9	0,95	A
E- 4	Las Tunas	MSc	Auxiliar	1	0,9	0,95	A
E- 5	Las Tunas	MSc	Auxiliar	1	0,9	0,95	A
E-6	Las Tunas	MSc	Auxiliar	1	1	1	A

E-7	Las Tunas	MSc	Auxiliar	0,9	0,85	0,85	A
E-8	Las Tunas	MSc	Auxiliar	0,9	0,85	0,85	A
E-9	Las Tunas	MSc	Auxiliar	1	0,9	0,95	A
E-10	Las Tunas	MSc	Auxiliar	1	0,9	0,95	A
E-11	Las Tunas	MSc	Auxiliar	1	1	1	A
E-12	Las Tunas	MSc	Auxiliar	0,9	0,85	0,85	A
E-13	Las Tunas	MSc	Auxiliar	1	0,9	0,95	A
E-14	Las Tunas	MSc	Auxiliar	0,9	0,85	0,85	A
E-15	Las Tunas	MSc	Auxiliar	1	0,9	0,95	A

⁺: Coeficiente de conocimiento

⁺⁺: Coeficiente de argumentación

⁺⁺⁺: Coeficiente de competencia

En la tabla 2 se muestra la evaluación del instrumento teórico por expertos según la escala de contenido propuesta, donde se utilizan el índice de correlación W de Kendall para la valoración de la significación de la concordancia entre los jueces; el criterio V de Aiken para evaluar la relevancia del contenido y la correlación de elementos corregida y el alfa de Cronbach para evaluar la consistencia entre la valoraciones de las calificaciones para validar el instrumento teórico.

Al examinar la concordancia entre las evaluaciones de los jueces. Acorde al estadístico W de Kendall, se aprecia que existe una concordancia de un 0.73 con un valor de $p=0.000$ menor que el nivel de significación $\alpha=0.05$.

El juicio valorativo sobre la relevancia del contenido de los ítems expresada mediante el criterio V de Aiken se estimó globalmente entre 0.63-0.84, exceptuando los ítems 2, 3, 6, 12 y 15 en lo que las estimación de dicho criterio oscilo entre 0-0.25.

Mientras que, la correlación ítem total se estimó por encima de 0.60 en todos los ítems alcanzando un máximo en los ítems 7 y 12 con 0.95, posibilitando que finalmente el alfa de Cronbach alcance una estimación de 0.96 para la totalidad el instrumento.

Tabla 2. Evaluación del instrumento teórico por expertos según la escala de contenido propuesta.

Ítem evaluado	Rango promedio	Media	Correlación		
			total de elementos corregida ⁺	V Aiken	IC 95%
1	11.17	3.87	0.74	0.75	0.63 - 0.84

2	3.07	1.67	0.65	0.25	0.16 - 0.37
3	4.33	2.07	0.91	0.25	0.16 - 0.37
4	8.67	3.27	0.69	0.75	0.63 - 0.84
5	8.33	3.07	0.91	0.75	0.63 - 0.84
6	3.37	1.53	0.88	0	0 - 0.06
7	9.5	3.47	0.95	0.75	0.63 - 0.84
8	9.93	3.53	0.88	0.75	0.63 - 0.84
9	9.83	3.53	0.83	0.75	0.63 - 0.84
10	11.17	3.87	0.85	0.75	0.63 - 0.84
11	10.73	3.73	0.88	0.75	0.63 - 0.84
12	4.57	2.2	0.95	0.25	0.16 - 0.37
13	11.17	3.87	0.60	0.75	0.63 - 0.84
14	11.47	3.93	0.74	0.75	0.63 - 0.84
15	2.7	1.47	0.94	0.25	0.16 - 0.37
Alfa de Cronbach					0.968
V Aiken Global				Puntual	0.75
				IC	0.63-0.84
W Kendall ⁺⁺					0.73
Chi Cuadrado 14gl					153.748
Significación asintótica					0.0000

⁺: procedente de hoja de estimación de la correlación total de elementos corregida.

⁺⁺: procedente de hoja de resultados test W de Kendall

gl: grados de libertad

En la tabla 3 se muestran los resultados del test W de Kendall, el índice de concordancia de V de Aiken, y la correlación de elementos corregida, para validar la escala de evaluación para la calificación del instrumento teórico.

La concordancia entre las evaluaciones de los jueces según al estadístico W de Kendall, se aprecia que existe una concordancia de un 0.71 con un valor de $p=0.000$ menor que el nivel de significación $\alpha=0.05$.

La relevancia del contenido de los ítems expresada mediante el criterio V de Aiken se estimó globalmente entre 0.63-0.84, exceptuando los ítems 4,7,9,10 y 13 en los que la estimación de dicho criterio oscilo entre 0-0.25.

En cambio, la correlación ítem total se estimó por encima de 0.50 en casi todos los ítems alcanzando un máximo en el Ítem 2 y 5 con 0.81 excepto el ítem 4 (referido a la exposición de argumentos conciliadores pertinentes a la selección de los mejores procederes) con un 33 %, el alfa de Cronbach alcanza una estimación de 0.87 para la totalidad el instrumento.

Tabla 3. Evaluación de la escala de evaluación del instrumento teórico por expertos.

Ítem evaluado	Rango promedio	Media	Correlación	V Aiken	IC 95%
			total de elementos corregida ⁺		
1	9.80	3.53	0.67	0.75	0.63 - 0.84
2	11.17	3.87	0.81	0.75	0.63 - 0.84
3	10.73	3.73	0.64	0.75	0.63 - 0.84
4	5.00	2.33	0.33	0.25	0.16 - 0.37
5	11.17	3.87	0.81	0.75	0.63 - 0.84
6	11.20	3.87	0.83	0.75	0.63 - 0.84
7	2.70	1.47	0.67	0.25	0.16 - 0.37
8	11.17	3.87	0.81	0.75	0.63 - 0.84
9	3.03	1.67	0.62	0.25	0.16 - 0.37
10	4.30	2.07	0.56	0.25	0.16 - 0.37
11	8.67	3.27	0.65	0.75	0.63 - 0.84
12	8.33	3.07	0.53	0.75	0.63 - 0.84
13	3.37	1.53	0.55	0.00	0 - 0.06
14	9.47	3.47	0.62	0.75	0.63 - 0.84
15	9.90	3.53	0.51	0.75	0.63 - 0.84
Alfa de Cronbach					0.869
V Aiken Global				Puntual	0.75
				IC	0.63-0.84
W Kendall ⁺⁺					0.712
Chi Cuadrado 14gl					153.748
Significación asintótica					0.0000

⁺: procedente de hoja de estimación de la correlación total de elementos corregida.

⁺⁺: procedente de hoja de resultados test W de Kendall

gl: grados de libertad

Discusión

En la confección de escalas son los expertos quienes determinan si, en su concepto, el instrumento en apariencia mide las cualidades deseadas; es importante mencionar que al momento de seleccionar a las personas que participarán en la evaluación del instrumento de medición/evaluación se debe prever que conozcan el tema a tratar, ya sea por experiencia en el campo laboral, por su formación profesional o por su trayectoria académica.¹⁶

Los resultados que se exponen en la tabla 1, se corresponden con la investigación de Hidalgo Anias y Márquez Pérez,¹⁷ relacionada con la validación de una escala de trastornos neuróticos, donde el 100 % de los expertos obtuvo un coeficiente de competencia alto (entre 0,94 y 1). Es deseable que, como sucedió en la presente investigación, los expertos que participen en todo proceso evaluativo tengan el máximo nivel de competencia, pues ello aporta rigor y oportunidades de mejora.

Herrera Masó y colaboradores,¹² exponen que el método de consulta a expertos propicia el contexto adecuado para obtener la mayor cantidad de información posible de los evaluadores consultados, atenuando el efecto de aspectos que no pueden ser controlados por el investigador y de otros factores relacionados con la complejidad, dificultad u obviedad del contenido sometido a su consideración.

En la investigación se logró la participación de tres profesionales pertenecientes a la Cátedra de Oclusión Dentaria “Plácido Ardanza Zulueta” de la provincia Camagüey y que implica al territorio donde se desplegó el presente estudio. Tal elemento que concuerda con lo expresado en su investigación Urrutia Egaña y colaboradores,¹⁸ al hacer referencia a Ortega Mohedano, quien asegura que la distribución de expertos en el área geográfica o la actividad laboral son factores importantes para la elección de los profesionales en función de experto.

Al analizar las sugerencias de los expertos, se concluyó que estas estuvieron relacionadas con la redacción científica o con el empleo de algunos términos específicos más que con la propia conceptualización u operacionalización como tal de las variables. Teniendo en cuenta, además, los resultados cuantitativos, no se consideró necesario reformular ninguna de las pautas para una segunda ronda de intercambio. De este modo, se demostró la validez de contenido del instrumento.

La concordancia obtenida en la evaluación del instrumento teórico, sugiere que los jueces, en sus evaluaciones, consideran la pertinencia de conservar cada uno de los ítems propuestos dentro del instrumento; las estimaciones de los estadígrafos expuestos para el cálculo de la confiabilidad del instrumentos teórico son similares a los expuestos por Bonilla Algovia y colaboradores,¹⁹ que obtienen cuantías superiores a 0.80, al viabilizar la validación de una escala de enfoque diferente

dirigida a la medición de las actitudes sexistas hostiles y benévolas en la población adolescente de Castilla-La Mancha.

La validación del instrumento por expertos realizada por Fleites Did y colaboradores,¹¹ en Villa Clara, Cuba, evidenció que el 100 % de las preguntas fueron razonables y comprensibles, sensibles a variaciones, con suposiciones básicas justificables, con componentes claramente definidos y derivables de datos factibles de obtener; en la validación por Delphi el 100 % de las preguntas fueron evaluadas con las categorías de muy adecuado y bastante adecuado, esta no viene acompañada de un estadístico de concordancia propia de escalas ordinales; lo que puede considerarse prescindible cuando los juicios valorativos concuerdan en un 100 % único.

En la investigación “Validez de contenido por juicio de expertos de un instrumento para medir conocimientos, actitudes y prácticas sobre el consumo de sal en la población peruana” de Torres Malca y colaboradores,¹³ utilizaron la V de Aiken para calcular el grado de acuerdo entre los jueces, donde presentaron valores V superiores a 0,80, además, de un coeficiente de variación menor a 25 %, por tanto, ninguno de los ítems fue eliminado, resultados que coinciden con la presente investigación. En otro estudio realizado con Agudelo y colaboradores²⁰ donde de igual forma se evalúa la validez aparente de la escala por juicio de expertos los resultados expuestos se aprecian por encima del 0.8 hasta 1.

La valoración general de los jueces para el instrumento consideró como pertinente la permanencia de los ítems sugeridos, interpretado como una medida de coherencia y validez.

Según lo mostrado en la tabla 3, la valoración general de los jueces sobre la calificación del instrumento y su escala, resultó pertinente. También consideraron conservar la calificación mediante la escala cualitativa nominal dicotómica, (1- Ausente, 2- Presente) del ítem dentro del instrumento.

Los resultados en cuanto a la concordancia interevaluador no se corresponden con investigaciones realizadas por Lorca y colaboradores²¹ para evaluar las propiedades psicométricas de la escala de estado funcional post-COVID-19. Se reveló que hubo concordancia entre los evaluadores solamente para el aspecto de relevancia (W de Kendall = 0,111; $p = 0,32$) en el resto de aspectos evaluados (suficiencia, claridad y coherencia) no se logra concordancia inter evaluador.

Otros estudios realizados que se hacen propicio mencionar debido a que exponen el diseño y la validación de un instrumento lo constituyen Tenango Leyte y Solís Martínez.²² En ellos se aplicó la prueba de Kendall, a partir de las 38 valoraciones de expertos, el coeficiente de concordancia $W = 0.498$ con una significancia $p < 0.000$, por lo que el panel de expertos estuvo de acuerdo con la evaluación de todos los ítems incluidos.

Conclusiones

El instrumento creado permitirá evaluar documentos científicos desde su contenido, unificar criterios y elevar el grado de preparación de los profesionales de la salud en la contemporaneidad.

Contribuciones de los autores

Conceptualización, MCSG, SMDG; metodología, MCSG, SMDG, KRVA; software, MCSG, KRVA; análisis formal, MCSG, SMDG, KRVA, EANO; escritura, revisión y edición, MCSG, SMDG, KRVA, EANO. Todos los autores han leído y aceptado la versión publicada del manuscrito.

Financiamiento

Esta investigación fue financiada por los propios autores.

Declaración de la junta de revisión institucional

El estudio fue realizado respetando lo establecido en la Declaración de Helsinki.

Declaración de consentimiento informado

Se obtuvo el consentimiento informado de todos los participantes involucrados en el estudio.

Declaración de disponibilidad de datos

Para obtener más información respecto a los datos puede comunicarse directamente con los autores de este estudio.

Agradecimientos

Los autores agradecen a cada uno de los expertos que participaron en la validación del cuestionario y tomaron un poco de su tiempo para contribuir con este valioso trabajo.

Conflictos de Interés

Los autores declaran no tener ningún conflicto de intereses.

Referencias

1. Cascaes da Silva F, Valdivia Arancibia BA, da Rosa Iop R, Barbosa Gutierrez Filho PJ, da Silva R. Escalas y listas de evaluación de la calidad de estudios científicos. Rev Cuba Inf Cienc Salud [Internet]. 2013 [citado el 17 de junio de 2024];24(3):295–312. Disponible en: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S2307-21132013000300007&lng=es.
2. Portocarrero Ramos CA, Bonifaz de Portocarrero RP. Diseño y Construcción de Escalas de Medición. Editorial Portocarrero Ramos, Carlos Alberto. [Internet] Primera edición digital. 2021. Disponible en: https://www.researchgate.net/publication/354322263_BNP_Diseño_y_construcción_de_escalas_de_medición
3. López Alonso SR, Morales Asencio JM. ¿Para qué se administran las escalas, cuestionarios, tests e índices? Index Enferm [Internet]. 2005 [citado el 17 de junio de 2024];14(48–49):7–8. Disponible en: https://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S1132-12962005000100001

4. Herrera Masó JR, Calero Ricardo JL, González Rangel MÁ, Collazo Ramos MI, Travieso González Y. El método de consulta a expertos en tres niveles de validación. Rev habanera cienc médicas [Internet]. 2022 [citado el 17 de junio de 2024];21(1). Disponible en: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1729-519X2022000100014&lng=es
5. Castellanos Serra L. Las Buenas Prácticas de Publicación, su evolución y el impacto esperado en salud pública. Rev Cub Salud Publica [Internet]. 2012 [citado el 17 de junio de 2024];38:725–33. Disponible en: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-34662012000500006&lng=es
6. Flores-Girón H, Paz C, Paz-Maldonado E. Explorando las competencias conceptuales básicas: Validez de contenido de un cuestionario sobre Radiología Oral y Maxilofacial. Bionatura [Internet]. 2023;8(3):1–6. Disponible en: <http://dx.doi.org/10.21931/rb/2023.08.03.42>
7. Muñiz J, Fonseca-Pedrero E. Diez pasos para la construcción de un test. Psicothema. 2019;31(1):7–16. Disponible en: <https://doi.org/10.7334/psicothema2018.291>
8. Gómez SMD, Miralles MD. Segundo Taller Nacional de Oclusión de la Cátedra Multidisciplinaria “Plácido Ardanza Zulueta”, 2017. Rev Cubana Estomatol [Internet]. 2018 [citado el 17 de junio de 2024];55(2):56–8. Disponible en: <http://www.revestomatologia.sld.cu/index.php/est/article/view/1807>
9. De Gante C, González S, Ortega B, Castillo E, Fernández S, Escala De Likert A. Una alternativa para elaborar e interpretar un instrumento de percepción social, 2020. Revista de la alta tecnología y sociedad. 2020;12(1). Disponible en: <file:///C:/Users/CUBA/Downloads/1.EscaladeLikert.pdf>
10. Cedeño C, Mena S. El Método Delphi Cualitativo y su Rigor Científico. Una revisión argumentativa. Revista Sociedad & tecnología. 2022;530-4. Disponible en: <https://dialnet.unirioja.es/servlet/articulo?codigo=8706211>
11. Pérez AMS, Piñero JP, Puig PL. Construcción y validación de un instrumento para evaluar el clima organizacional en el ámbito de la salud pública. Rev Cub Salud Publica [Internet]. 2015 [citado el 18 de junio de 2024];41(4). Disponible en: <https://revsaludpublica.sld.cu/index.php/spu/article/view/513>
12. Moriyama IM. Indicators of social change. Problems in the measurements of health status. New York: Russell Sage Foundation; 1968.p.593. Disponible en: <http://scielo.sld.cu/scieloOrg/php/reflinks.php?refpid=S0864-2125200900040000900005&lng=pt&pid=S0864-21252009000400009>
13. Torres-Malca JR, Vera-Ponce VJ, Zuzunaga-Montoya FE, Talavera JE, De La Cruz-Vargas JA. Validez de contenido por juicio pericial de un instrumento para medir conocimientos, actitudes y

prácticas sobre el consumo de sal en la población peruana. Rev Fac Med Humana [Internet]. 2022;22(2):273-9. Disponible en: <http://www.scielo.org.pe/pdf/rfmh/v22n2/2308-0531-rfmh-22-02-27314>.

14. Frías-Navarro D. Apuntes de estimación de la fiabilidad de consistencia interna de los ítems de un instrumento de medida. Universidad de Valencia. España. 2022. Disponible en: <https://www.uv.es/friasnav/AlfaCronbach>

15. Barrueta Gómez N, Peña Martínez S, Fernández Sánchez E. El estadígrafo Kendall y su aplicación. Un ejemplo práctico. Editorial A3Manos.La Habana, Cuba; [Internet].2022.vol. 9, núm. 16. Disponible en: <http://portal.amelica.org/ameli/journal/784/7843889004/html/>

16. Rodríguez Medina MA, Poblano-Ojinaga ER, Alvarado Tarango L, González Torres A, Rodríguez Borbón MI. Validación por juicio de expertos de un instrumento de evaluación para evidencias de aprendizaje conceptual. RIDE Rev Iberoam Para Investig Desarro Educ [Internet]. 2021 [citado el 18 de junio de 2024];11(22). Disponible en: <https://www.ride.org.mx/index.php/RIDE/article/view/960>

17. Hidalgo Anias NM, Márquez Pérez JF. Validación de una escala de trastornos neuróticos. Humanid médicas [Internet]. 2012 [citado el 18 de junio de 2024];12(1):9–21. Disponible en: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1727-81202012000100002&lng=es

18. Urrutia Egaña M, Barrios Araya S, Gutiérrez Núñez M, Mayorga Camus M. Métodos óptimos para determinar validez de contenido. Educ médica super (Impresa) [Internet]. 2014 [citado el 18 de junio de 2024];28(3):547–58. Disponible en: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-21412014000300014&lng=es

19. Bonilla-Algovia E, Ibáñez Carrasco M, Carrasco Carpio C. Validez de la Escala de Detección de Sexismo en Adolescentes (DSA) en Castilla-La Mancha, España. Psykhe (Santiago) [Internet]. 2022;31(2). Disponible en: <http://dx.doi.org/10.7764/psykhe.2021.28041>

20. Agudelo-Hernández F, Vélez-Botero H, Rojas-Andrade R. Traducción y adaptación de la Escala de Continuidad de Servicios de Salud Mental de Alberta en un contexto latinoamericano. Rev Chil Neuro-Psiquiatr [Internet]. 2023 [citado el 18 de junio de 2024];61(3):258–69. Disponible en: http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0717-92272023000300258&lng=es

21. Lorca LA, Leão Ribeiro I, Torres-Castro R, Sacomori C, Rivera C. Propiedades psicométricas de la escala Post-COVID-19 Functional Status para adultos sobrevivientes de COVID-19. Rehabil (Madr, Internet) [Internet]. 2022;56(4):337–43. Disponible en: <http://dx.doi.org/10.1016/j.rh.2021.07.002>

22. Leyte CT, Martínez ES. Diseño y validación de un instrumento de evaluación del compromiso docente en educación media superior. *Ciencia Latina* [Internet]. 2022 [citado el 18 de junio de 2024];6(4):1201–23. Disponible en: <https://ciencialatina.org/index.php/cienciala/article/view/2656>.

| **Received:** 18 June 2024 | **Accepted:** 15 August 2024 | **Published:** 15 September 2024 |

Citation: Suárez-González, M., Díaz-Gómez, S., Vazquez-Argote, K., Núñez-Oduardo, E. Escala para documentos científicos de estandarización médica. *Bionatura*. 2024; Volume (9). No 3.

Peer review information: Bionatura thanks the anonymous reviewers for their contribution to the peer review of this work using <https://reviewerlocator.webofscience.com/>

All articles published by Bionatura Journal are freely and permanently accessible online immediately after publication, without subscription charges or registration barriers.

Publisher's Note: Bionatura stays neutral concerning jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2024 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>)