

<https://doi.org/10.23913/ride.v12i24.1183>

Artículos científicos

Análisis de validez de contenido por criterio de jueces de un instrumento para evaluar un manuscrito

Content validity analysis by judge's criteria of an instrument to evaluate a manuscript

Análise da validade de conteúdo por critérios de juízes de um instrumento de avaliação de um manuscrito

Rigoberto Reyes Valenzuela

Tecnológico Nacional de México/ IT Orizaba, México

rigoreyes23@hotmail.com

<https://orcid.org/0000-0002-8570-8563>

Domingo Noé Marrón Ramos

Tecnológico Nacional de México/ IT Milpa Alta, México

dmarron22@hotmail.com

<https://orcid.org/0000-0003-1964-6592>

Arturo González Torres

Tecnológico Nacional de México/ IT Milpa Alta, México

cann.azteca13@gmail.com

<https://orcid.org/0000-0002-3337-7600>

Rubén Juárez Rodríguez

Tecnológico Nacional de México/ IT Orizaba, México

Ru_juod@hotmail.com

<https://orcid.org/0000-00023-0002-1787>

Fátima Yaraset Mendoza Montero

Tecnológico Nacional de México/ IT Milpa Alta, México

fatmendez71@hotmail.com

<https://orcid.org/0000-0003-0240-5434>



Resumen

Ante los altos niveles de exigencia de las revistas científicas, las cuales requieren una gestión editorial exitosa, es necesario tener herramientas que sirvan para obtener valoraciones confiables que ofrezcan una ventaja competitiva dentro del ámbito de la investigación. Por eso, el propósito de esta investigación fue examinar la validez de contenido a través del juicio de árbitros para evaluar un manuscrito. Esta investigación se realizó en tres etapas: la primera consistió en las bases teóricas. La segunda comprendió el proceso de diseño. La tercera abarcó la validez. Esta última fue elaborada a través de un grupo de ocho jueces y para la validez de criterio se utilizó la prueba Friedman. Los hallazgos de la presente investigación muestran un acuerdo significativo entre los expertos para la validez del contenido para el manuscrito evaluado.

Palabras clave: criterio, instrumento, jueces, validez.

Abstract

Given the high levels of demand in scientific journals that dictate having a successful editorial management, it is necessary to have tools that serve to obtain reliable evaluations, which will provide a competitive advantage within the field of research. The purpose of this study was to analyze the content validity by judging criteria of an instrument to evaluate a manuscript. This research was carried out in three stages: the first consisted of the theoretical bases. The second, he understood the design process. The third covered validity. The content validity was elaborated through a panel of eight judges and the Friedman test was used for the criterion validity. The findings of the present investigation show a significant agreement among the judges for the categories of clarity, coherence, relevance and sufficiency.

Keywords: criteria, instrument, judges, validity.

Resumo

Diante dos altos níveis de demanda dos periódicos científicos, que exigem uma gestão editorial bem-sucedida, é necessário contar com ferramentas que sirvam para obter avaliações confiáveis que ofereçam uma vantagem competitiva dentro do campo de pesquisa. Portanto, o objetivo desta pesquisa foi examinar a validade de conteúdo por meio do julgamento de pareceristas para avaliar um manuscrito. Esta pesquisa foi realizada em três etapas: a primeira consistiu nas bases teóricas. A segunda compreendeu o processo de design. A terceira abrangia a validade. Este

último foi elaborado por meio de um grupo de oito juízes e foi utilizado o teste de Friedman para validade de critério. Os achados da presente investigação mostram uma concordância significativa entre os especialistas quanto à validade de conteúdo do manuscrito avaliado.

Palabras-clave: critério, instrumento, juízes, validade.

Fecha Recepción: Septiembre 2021

Fecha Aceptación: Febrero 2022

Introducción

Actualmente, existe un creciente interés por parte de los responsables de las revistas científicas para llevar a cabo periódicamente procesos de evaluación de esos medios de divulgación con el fin de promover su calidad (Borrego y Urbano, 2006). En concordancia con esta idea, desde el año 2020 se empezó a trabajar en la mejora de los procesos editoriales de una revista arbitrada e indexada con reconocimiento internacional con el fin de garantizar la calidad de sus manuscritos, los cuales son producidos por investigadores, profesores y estudiantes de diversos campos de la ingeniería.

Según Vildózola (2003), para alcanzar una gestión editorial de prestigio, los manuscritos recibidos deben pasar por un proceso de evaluación enfocado en la estructura, redacción y calidad científica de dichos textos, para lo cual es importante contar con instrumentos que faciliten su proceso de evaluación. En tal sentido, Elosua (2003) considera que el propósito de las investigaciones centradas en el tema de validez debe ser recolectar evidencias suficientes para la correcta toma de decisión en la interpretación de los datos analizados.

Salas (1998) argumenta que el significado de validez corresponde a lo que se quiere medir por parte de un instrumento contra lo que realmente se está evaluando, lo cual debe ser comparado para verificar que se ha cumplido con las exigencias solicitadas. Así pues, la validez —según García (2002)— se asocia con la consistencia de las situaciones necesarias en la que un acto sucede en su vigencia y autenticidad. La Asociación Estadounidense de Investigación Educativa —también conocida como AERA—señala que, de acuerdo con sus estándares, la validez es referida al nivel de congruencia que existe con la conjetura que apoya a la propia herramienta de medición y la certidumbre experimental obtenida en su implementación. Aiken (2003), sin embargo, objeta que el significado de autenticidad de un tema es referido a un banco de datos, los cuales son medidos por un instrumento que debe brindar un resultado confiable. Rodríguez y Sossa (2004) mencionan que la validez se divide de este modo: validez de contenido, validez de criterio, validez de apariencia y validez de constructo. Para Piratoba (2011) la fiabilidad junto con la validez son los criterios que fijan la calidad de un instrumento.

Por tal motivo, la presente investigación tiene como objetivo analizar la validez de contenido por criterio de jueces de un instrumento para evaluar un manuscrito, para lo cual se empleó la selección de expertos. En tal sentido, se aplicó la norma NC 49:1981 con el fin de conocer el número de árbitros necesarios para cumplir con el estudio, y se aplicó la técnica Delphi, adecuada para el contraste de la hipótesis (Ortega, 2008).

Asimismo, se usó la herramienta gratuita Formulario de Google, que sirve para obtener mayores beneficios en cuanto a su empleo (Lorenzo *et al.*, 2017). Para la valoración de los resultados de los jueces se utilizó la prueba no paramétrica de Friedman debido a que es la que se recomienda cuando se comparan tres o más muestras relacionadas (Berlanga y Rubio, 2012). Los resultados obtenidos aportan información consistente de la existencia de un acuerdo significativo entre los expertos.

Metodología

Participantes

Para el estudio se seleccionó de forma no probabilística y discrecional a las personas que fungieron como jueces, las cuales cuentan con el conocimiento y la experiencia en el tema. Se empleó la norma cubana NC 49:1981 para determinar el número de expertos.

En la investigación se estableció como vía fundamental para la comunicación el correo electrónico, y para la contestación del instrumento se empleó la herramienta Formularios de Google. Durante el transcurso del estudio se mantuvo contacto con cada uno de los jueces para responder sus dudas y comentarios. Los integrantes seleccionados forman parte del comité de la revista, lo cual facilitó el propósito de la investigación; además, se adquirió una discusión colectiva con incidencia directa o indirecta en la objetividad de sus juicios.

Método

En el presente estudio se empleó un enfoque cuantitativo, no experimental, transversal y descriptivo. Cuantitativo porque se partió de una problemática que, para desarrollarla y culminarla con éxito, requería de un orden en su ejecución. Descriptivo porque se buscó conocer la problemática estudiada y el comportamiento de las variables inmiscuidas exponiendo los resultados recabados. Transversal porque en un tiempo determinado se analizó, por medio de instrumentos de evaluación, la problemática de una manera real, y no experimental porque las variables involucradas en el estudio no fueron manipuladas.

Procedimiento

La investigación se desarrolló en tres fases sugeridas por Poblano *et al.* (2019), García *et al.* (2021) y Rodríguez *et al.* (2021). En la primera se elaboraron las bases teóricas, para lo cual se inició con el planteamiento del estudio. Posteriormente, se desarrollaron el alcance y la justificación; además, se plantearon el propósito y el problema, los cuales en conjunto guiaron la presente investigación. En esta etapa se buscaron antecedentes con el fin de conocer la situación actual del tópico en la literatura existente. La información adquirida se consiguió en repositorios confiables como Redalyc, Scielo, Latindex, Google Académico, etc. Los criterios de selección fueron los siguientes: 1) el material debía incluir el tópico del estudio indagado, y 2) se revisaron libros académicos, así como artículos escritos en español e inglés publicados en revistas arbitradas e indexadas.

En la segunda etapa se procedió a desarrollar la validez, lo cual se efectuó con el apoyo de árbitros a quienes se les invitó a participar en la investigación. Para esto se aplicó el método Delphi. En concreto, se contó con la intervención de ocho jueces, los cuales forman parte de la revista.

En la tercera fase se implementaron las pruebas de validez de contenido con los árbitros seleccionados. Para que el estudio fuera objetivo, a cada juez se le garantizó el apoyo y la privacidad del estudio. Culminado el estudio y recolectada toda la información, se procedió al cálculo de la validez de contenido, para lo cual se aplicó el parámetro Friedman con el fin de determinar las diferencias en la ubicación central (mediana) para el análisis de ensayos con medidas repetidas unidireccionales que tienen tres o más muestras dependientes. Se plantearon las siguientes hipótesis para poder establecer el valor p , el cual ayudó a fijar el grado de acuerdo entre las personas (expertos) que intervinieron en el estudio. Las hipótesis fueron:

$$H_0: X \sim N(\mu, \sigma^2)$$

$$H_1: X \not\sim N(\mu, \sigma^2)$$

Resultados

La primera parte consistió en la identificación de las variables que debía contener la evaluación de un manuscrito, por lo que se llevó a cabo una revisión teórica de varios instrumentos de evaluación de contribuciones en México y en otras partes del mundo. Se realizó un consenso de los criterios encontrados para después solicitar a tres integrantes de la revista la valoración de las variables localizadas en la revisión de literatura con el fin de que emitieran su

opinión y sugerencias. El resultado de esta fase fue obtener el primer borrador del instrumento de evaluación, que incluyó cuatro criterios y dieciocho condiciones de calidad (tabla 1).

Tabla 1. Lista de criterios de calidad para el instrumento de evaluación

CONSTRUCTOS	ASPECTOS POR EVALUAR
ENCABEZADO	Título Autor (es) Institución Correo electrónico
RESUMEN	En español Palabras clave En inglés Keywords
ESTRUCTURA	Introducción Desarrollo Análisis de resultados Conclusiones y discusión Referencias Extensión
EVALUACIÓN	Dimensión Recomendación Comentarios adicionales Identificación del evaluador

Fuente: Elaboración propia

En la fase del diseño del instrumento de evaluación se tomaron en cuenta las recomendaciones de los integrantes de la revista. Las sugerencias fueron de complemento para algunos criterios preestablecidos, los cuales se volvieron a actualizar, así como de cambios en algunas oraciones para facilitar su comprensión (tabla 2).

Tabla 2. Instrumento para evaluación de manuscritos por juicio de expertos

	Aspecto por evaluar	Descripción
Encabezado	Título	El título se encuentra escrito en idioma español y contiene en su extensión de 16 a 18 palabras claves.
	Autor	Se señala la formación académica y el nombre completo de la persona.
	Institución	Se identifica la dirección completa de la institución de procedencia de la persona.
	Correo electrónico	Se estipula la cuenta de correo de la persona.
Resumen	En español	Responde a las preguntas ¿qué se hizo?, ¿cómo se hizo? y ¿a qué resultados llegó? Contiene una extensión máxima de 350 palabras.
	Palabras claves	Contiene de 3 a 7 palabras escritas en idioma español.
	En inglés (<i>abstract</i>)	Se encuentra traducido al idioma inglés.
	En inglés (<i>key words</i>)	Se encuentran traducidas al idioma inglés.
Estructura	Introducción	Se señalan las referencias del tópico y el problema de estudio. Se estipula el propósito del trabajo con énfasis en la justificación de la contribución presentada.
	Desarrollo	Se relata con entendimiento la metodología manejada; además, se incluyen cálculos que sustenten la contribución presentada.
	Análisis de resultados	Se expresa fácilmente y con una lógica los resultados logrados; además, estos responden al propósito de la investigación.
	Conclusiones y discusión	Se pronuncia claramente y se discuten las evidencias con otros estudios; además, se adhieren datos para futuras investigaciones.
	Referencias	Se presentan extraídas de fuentes confiables, se contabilizan al menos 20 y se ajustan a la norma Vancouver.
	Extensión	La conformación de la contribución cumple de 10 a 12 páginas, teniendo en cuenta un total máximo de 5 figuras y 5 tablas.
Evaluación	Dimensión	La contribución posee una lógica en su contenido, es novedosa y aporta al ámbito de la disciplina.
	Recomendación	La contribución debe de aceptarse como se presenta, debe de tener correcciones significativas, debe de poseer una revisión estricta o debe de no aceptarse.
	Comentarios adicionales	La contribución debe de incluir alguna recomendación por usted. De ser afirmativa su respuesta, por favor estipular su opinión.

	Identificación del juez	Presenta el visto bueno y el nombre del árbitro.
--	-------------------------	--

Fuente: Elaboración propia

En la etapa de validación del instrumento de evaluación por expertos, primero se determinó la cantidad de personas que deberían participar en el estudio. Para establecer dicho número Lao, Pérez y Marrero (2016) recomiendan seguir el procedimiento aprobado por la Oficina Nacional de Normalización, y estipulada por la NC 49:1981, la cual refiere que, para lograr un nivel de confianza y calificación elevada, el grupo de expertos debe oscilar entre 7 y 15. Para el estudio se empleó una proporción estimada de error (P) de 0.05, un nivel de confianza (K) de 95 % y un nivel de precisión (i) de 0.15. El resultado arrojado fue de 8 expertos para la validación de contenido de la presente investigación.

La consulta se realizó a 11 personas que forman parte del equipo de la revista. Una vez completado el estudio, 8 jueces lograron un coeficiente de competencia aprobatorio, dado que sus resultados se ubicaron entre medio y alto, lo cual —según Cruz (2006) y Hernández y Robaina (2017)— los hace formar parte del equipo de expertos para evaluar el instrumento de esta investigación. La tabla 3 representa el resultado final del método Delphi.

Tabla 3. Resultados del procesamiento para la determinación del coeficiente de competencia de los expertos

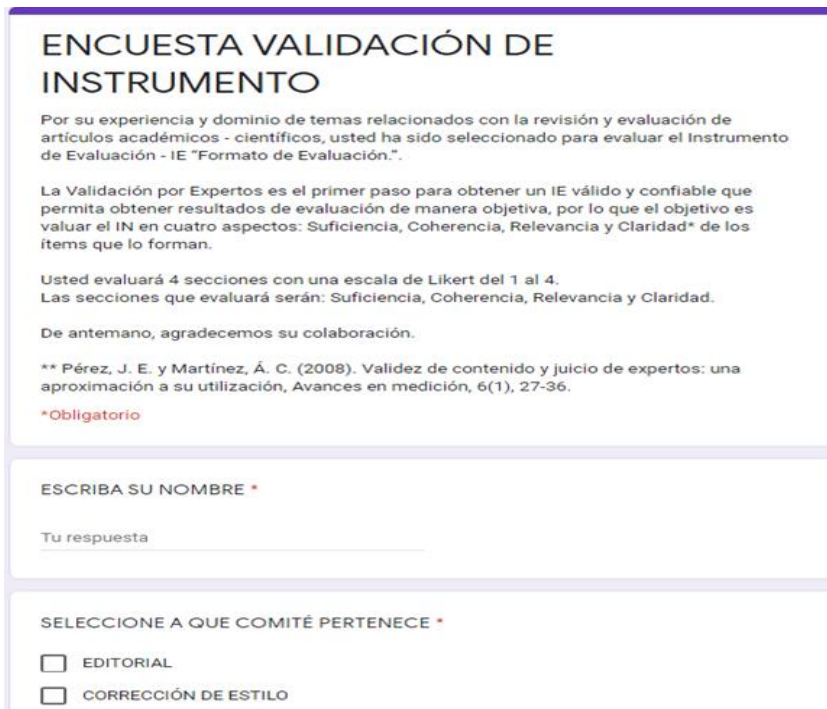
RESULTADO FINAL MÉTODO DELPHI					
$K = 0.5 \times (K_c + K_a)$					
EXPERTO	Ka	Kc	K	Resultado	Decisión
1	0.90	1.05	0.975	Alta valoración	Sí
2	0.70	0.95	0.825	Alta valoración	Sí
3	0.80	0.65	0.725	Media valoración	No
4	0.90	0.95	0.925	Alta valoración	Sí
5	0.50	0.95	0.725	Media valoración	No
6	0.80	0.95	0.875	Alta valoración	Sí
7	0.50	0.85	0.675	Baja valoración	No
8	0.80	1.05	0.925	Alta valoración	Sí
9	0.90	1.05	0.975	Alta valoración	Sí
10	0.80	0.95	0.875	Alta valoración	Sí
11	0.90	0.95	0.925	Alta valoración	Sí

Fuente: Elaboración propia con base en Cruz (2006) y Hernández y Robaina (2017)

Una vez analizado los resultados de la tabla anterior, se solicitó al grupo de 8 expertos que evaluaran la lista de cotejo en cuatro diferentes áreas: suficiencia, relevancia, claridad y

coherencia (Escobar y Cuervo, 2008). La figura 1 representa la encuesta que se envió vía Formulario de Google.

Figura 1. Formulario para evaluación de manuscritos por juicio de expertos



ENCUESTA VALIDACIÓN DE INSTRUMENTO

Por su experiencia y dominio de temas relacionados con la revisión y evaluación de artículos académicos - científicos, usted ha sido seleccionado para evaluar el Instrumento de Evaluación - IE "Formato de Evaluación."

La Validación por Expertos es el primer paso para obtener un IE válido y confiable que permita obtener resultados de evaluación de manera objetiva, por lo que el objetivo es evaluar el IN en cuatro aspectos: Suficiencia, Coherencia, Relevancia y Claridad* de los items que lo forman.

Usted evaluará 4 secciones con una escala de Likert del 1 al 4.
Las secciones que evaluará serán: Suficiencia, Coherencia, Relevancia y Claridad.

De antemano, agradecemos su colaboración.

** Pérez, J. E. y Martínez, Á. C. (2008). Validez de contenido y juicio de expertos: una aproximación a su utilización, *Avances en medición*, 6(1), 27-36.

*Obligatorio

ESCRIBA SU NOMBRE *

Tu respuesta

SELECCIONE A QUE COMITÉ PERTENECE *

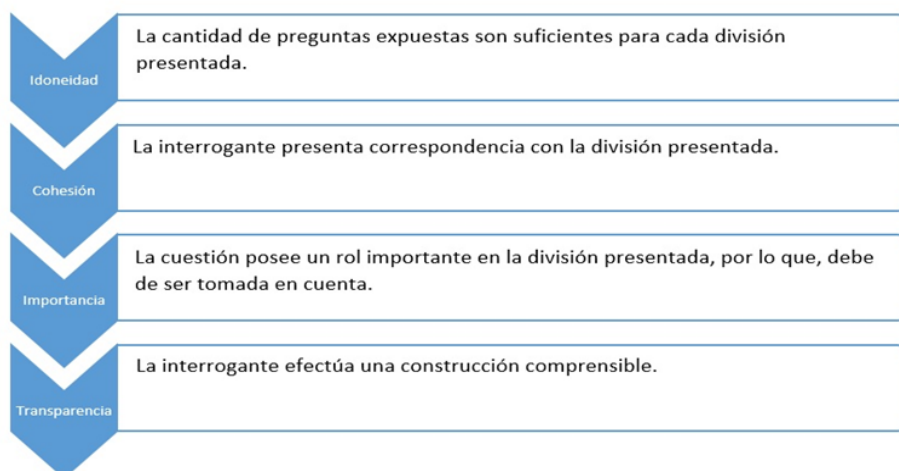
EDITORIAL

CORRECCIÓN DE ESTILO

Fuente: Elaboración propia

La técnica Delphi —según van Der *et al.* (2002)— ayuda a obtener una sinergia entre los participantes. La figura 2 representa información suministrada por Escobar y Cuervo (2008), la cual está conformada por clases y criterios utilizados para valorar un instrumento.

Figura 2. Categorías empleadas para la valoración del instrumento



Fuente: Escobar y Cuervo (2008, pp. 35-36)

A manera de ejemplo, en relación con el criterio coherencia, los expertos calificaron a cada ítem con un valor (entre 1 y 4) según consideraron correcto o conveniente (tabla 4).

Tabla 4. Resultados de la evaluación por expertos al criterio *coherencia*

CRITERIO	EXPERTOS	ÍTEMS																
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
COHERENCIA	1	3	3	4	3	3	3	4	3	4	3	4	3	4	3	4	3	4
	2	3	2	4	4	3	2	4	4	3	2	4	4	3	2	4	4	4
	3	3	4	2	4	3	4	2	4	3	4	2	4	3	4	2	4	3
	4	3	3	3	3	3	3	3	3	2	3	3	3	4	3	3	3	4
	5	3	4	2	4	3	4	2	4	3	4	2	4	3	4	2	4	4
	6	3	4	3	4	3	4	3	4	3	4	3	4	3	4	3	4	3
	7	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	2
	8	2	2	3	2	2	2	3	2	2	2	3	2	2	2	3	2	3

Fuente: Elaboración propia

Posteriormente, los datos de la tabla 4 se analizaron estadísticamente con la ayuda del *software* SPSS (Statistical Package for Social Science), versión 22, para medir la distribución que siguen los datos cuantitativos, la cual Flores, Miranda y Villasís (2017) señalan que debe analizarse para precisar si poseen un comportamiento normal. Para ello, se aplicó la prueba de normalidad, la cual es recomendable para determinar si existe una distribución normal o no (Argote, 2009).

La tabla 5 muestra los resultados arrojados por el *software* SPSS. Para ello, se siguió la recomendación de Galindo (2020), quien señala que cuando se tiene una muestra mayor de 50 datos, la prueba que se debe emplear es la Kolmogorov Smirnov. Las hipótesis formuladas fueron las siguientes:

H₀: La muestra no simboliza discrepancia significativa

H₁: La muestra sí simboliza discrepancia significativa

Tabla 5. Resultados de prueba de normalidad para el criterio *coherencia*

	Kolmogorov-Smirnov ^a		
	Estadístico	Gl	Sig.
Coherencia	.248	272	.000

Fuente: Elaboración propia

El resultado mostrado en la tabla 5, en la columna *Sig.* (es decir, 0.000) es inferior a 0.05, por lo que se tiene que rechazar la hipótesis H₀; en consecuencia, se debe elegir la hipótesis nula (H₁), lo que significa que la muestra de datos no sigue una distribución normal. La tabla 6 representa los valores para los tres criterios restantes: suficiencia, claridad y relevancia.

Tabla 6. Resultados de prueba de normalidad para los criterios *suficiencia, claridad y relevancia*

	Kolmogorov-Smirnov ^a		
	Estadístico	Gl	Sig.
Suficiencia	.263	272	.000
Claridad	.302	272	.000
Relevancia	.304	272	.000

Fuente: Elaboración propia

Los efectos conseguidos de los tres criterios por medio del *software* SPSS dan como resultados un valor menor a 0.05; por tal motivo, para cada criterio se rechaza la hipótesis H_0 y se selecciona la hipótesis H_1 , con lo cual se puede afirmar que la distribución no sigue una distribución normal.

Para aplicar la prueba de Friedman en el criterio *coherencia*, el nivel de significancia que se manejó fue 0.05, el cual —según Juárez (2015)— es el mínimo aceptable establecido en ciencias sociales y en psicología. Las hipótesis planteadas fueron las sugeridas por Ramírez, Murcia y Castro (2014):

$$H_0: X \sim N(\mu, \sigma^2)$$

$$H_1: X \not\sim N(\mu, \sigma^2)$$

De la misma manera, se procedió a llevar a cabo la prueba de Friedman para los criterios *suficiencia, claridad y relevancia*. En la tabla 7 se exhiben los efectos emanados para cada criterio.

Tabla 7. Resultados de prueba Friedman para los criterios *coherencia, suficiencia, claridad y relevancia*

	Criterios			
	Coherencia	Suficiencia	Claridad	Relevancia
N	8	8	8	8
Chi-cuadrado	20.296	25.621	26.022	22.257
gl	33	33	33	33
Sig. asintótica	.959	.817	.801	.922

Fuente: Elaboración propia

Los resultados enseñados en la tabla 6 —en el renglón *Sig. Asintótica*— muestran valores mayores que 0.05; por lo tanto, se sigue la recomendación de Guisande, Vaamonde y Barreiro (2013), según la cual no existe certeza bastante para no objetar la hipótesis nula (H_0), y se asevera que, entre las opiniones de los árbitros, existe un acuerdo significativo en los cuatro criterios.

Discusión

La metodología Delphi resultó ser una técnica muy útil en el estudio de un consenso de árbitros, resultado similar a lo reportado por Wright *et al.* (1996), León y Montero (2003) y Calabuig y Crespo (2009). Asimismo, el análisis estadístico de la presente investigación en sus cuatro criterios para la validez de contenido (*coherencia, suficiencia, claridad y pertinencia*), para cada uno de ellos el estadístico de prueba *Sig. asintótica* tuvo un valor mayor a 0.050, y soporta la afirmación de Ramírez y Polack (2020), quienes mencionan que no existe evidencia suficiente para rechazar H_0 , dado que el valor *Sig. asintótica* es mayor que el nivel de alfa de 0.050. Esta afirmación puede sustentarse por lo descrito por los autores Quispe *et al.* (2019) quienes sugieren comparar los resultados de chi cuadrada y la chi crítica. Para la toma de decisión de no rechazar H_0 , la chi crítica debe ser mayor a la chi cuadrada. En la tabla 8 se presentan los resultados obtenidos.

Tabla 8. Características de la muestra

	Criterios			
	Coherencia	Suficiencia	Claridad	Relevancia
N	8	8	8	8
Chi cuadrada	20.296	25.621	26.022	22.257
gl	33	33	33	33
Chi crítica	47.339	47.339	47.339	47.339

Fuente: Elaboración propia

Las derivaciones logradas de las chi cuadradas y de las chi críticas señalan que las cuatro primeras son mayores que las segundas; en otras palabras, se concluye que existe concordancia entre los expertos sobre los elementos que debe evaluar el instrumento para la revisión de un artículo. Emplear el criterio de jueces en investigaciones que tengan como propósito la validación de contenido para instrumentos de evaluación ofrece una alternativa de utilización prudente en los recursos involucrados para la ejecución de los estudios.

Conclusiones

Como se observó en la presente investigación, la validación de contenido por medio del juicio de expertos exige un proceso sistemático, el cual en cada etapa involucrada debe realizarse por completo para obtener un resultado exitoso.

El propósito del presente estudio fue analizar la validez de contenido por criterio de jueces de un instrumento para evaluar un artículo de investigación. La herramienta que se construyó para tal efecto comprendió cuatro factores: encabezado, resumen, estructura y evaluación.

Los resultados señalan que la prueba Friedman valida el consenso de las opiniones en las cuatro variables evaluadas por los árbitros, lo que brinda un instrumento robusto en su contenido.

Para concluir, se sugiere desarrollar una investigación para ponderar las opiniones de los jueces. Para ello, será importante llevar a cabo cada etapa de una investigación de validez de contenido empleando el juicio de expertos en su totalidad con el fin de lograr un resultado confiable.

Futuras líneas de investigación

El presente estudio representa el primer paso de la validación de una herramienta para contribuir a la valoración de un manuscrito. En virtud de ello, la siguiente fase se enfocará en completar la validez de criterio y de constructo de la herramienta propuesta en esta investigación.

Una línea de investigación adicional será la de llevar a cabo un estudio de la confiabilidad (consistencia interna) del instrumento a través del coeficiente alfa de Conbrach por medio de una prueba o corrida piloto de una investigación en diferentes grupos.

Referencias

- Aiken, L. R. (2003). *Tests psicológicos y evaluación*. Pearson Educación.
- Argote, M. L. (2009). Comparación y evaluación de la distribución del estimador de la tasa global de fecundidad de Bolivia. *Papeles de Población*, 15(62), 201-222. Recuperado de <https://www.redalyc.org/articulo.oa?id=11212354007>
- Berlanga, V. y Rubio, M. J. (2012). Clasificación de pruebas no paramétricas. Cómo aplicarlas en SPSS. *REIRE: Revista de Innovación e Investigación en Educación*, 5(2), 101-113. Doi: <https://doi.org/10.1344/reire2012.5.2528>
- Borrego, Á. y Urbano, C. (2006). La evaluación de revistas científicas en ciencias sociales y humanidades. *Información, Cultura y Sociedad*, (14), 11-27. Doi: <https://doi.org/10.34096/ics.i14.886>
- Calabuig, F. y Crespo, J. (2009). Uso del método Delphi para la elaboración de una medida de la calidad percibida de los espectadores de eventos deportivos. *RETOS*, (15), 21-25. Doi: <https://doi.org/10.47197/retos.v0i15.34993>
- Cruz, M. (2006). *El método Delphi en las investigaciones educativas*. Academia.
- Escobar, J. y Cuervo, A. (2008). Validez de contenido y juicio de expertos: una aproximación a su utilización. *Avances en Medición*, (6), 27-36. Recuperado de https://www.humanas.unal.edu.co/lab_psicometria/application/files/9416/0463/3548/Vol_6_Articulo3_Juicio_de_expertos_27-36.pdf
- Elosua, P. (2003). Sobre la validez de los tests. *Psicothema*, 15(2), 315-321. Recuperado de <http://www.psicothema.com/pdf/1063.pdf>
- Flores, E., Miranda, M. G. y Villasís, M. Á. (2017). El protocolo de investigación VI: cómo elegir la prueba estadística adecuada. Estadística inferencial. *Revista Alergia México*, 64(3), 364-370. Doi: <https://doi.org/10.29262/ram.v64i3.304>
- Galindo, H. (2020). *Estadística para no estadísticos. Una guía básica sobre la metodología cuantitativa de trabajos académicos*. Ciencias Editorial.
- García, R., Poblano, E. R., Reyes, R., Cuamea, G. y Juárez, R. (2021). Elección de carrera e institución de educación superior: validación de instrumento de medición mediante el modelado de ecuaciones estructurales. *RIDE: Revista Iberoamericana para la Investigación y el Desarrollo Educativo*, 11(22). Doi: <https://doi.org/10.23913/ride.v11i22.961>

- García, S. (2002). La validez y la confiabilidad en la evaluación del aprendizaje desde la perspectiva hermenéutica. *Revista de Pedagogía*, 23(67), 297-318. Recuperado de http://ve.scielo.org/scielo.php?script=sci_arttext&pid=S0798-97922002000200006&lng=es&nrm=iso&tlng=es
- Guisande, C., Vaamonde, A. y Barreiro, A. (2013). *Tratamiento de datos con R, STATISTICA y SPSS*. Díaz de Santos.
- Juárez, C. S. (2015). *Análisis de varianza de tipo I: Anova I. Problemario de estadística aplicada* [material educativo]. Universidad Autónoma del Estado de México. Repositorio de la Universidad Autónoma del Estado de México. Recuperado de <http://ri.uaemex.mx/bitstream/handle/20.500.11799/33860/secme-16960.pdf?sequence=1&isAllowed=y>
- Hernández F. y Robaina J. I. (2017). Guía para la utilización de la metodología Delphi en las etapas de comprobación de productos terminados tipo software educativo. *Revista 16 de Abril*, 56(263), 26-31. Recuperado de http://www.rev16deabril.sld.cu/index.php/16_04/article/view/429/220
- Lao, Y. O., Pérez, M. C. y Marrero, F. (2016). Procedimiento para la selección de la comunidad de expertos con técnicas multicriterio. *Ciencias Holguín*, 22(1), 34-49. Recuperado de <https://www.redalyc.org/pdf/1815/181543577003.pdf>
- León, O. G. y Montero, I. (2003). *Métodos de investigación en psicología y educación* (3.ª ed.). Madrid: McGraw-Hill.
- Lorenzo, G., Lledó, A., Arráez, G., Lorenzo, A., González, C., Gómez, M., Hernández, M. J., Sanmartín, R., Urrea, M. E. y Vicent, M. (2017). Innovaciones en evaluación: Google Forms como herramienta de evaluación y retroalimentación de los aprendizajes del alumnado. En C. Roig (coord.), *Memorias del programa de redes-I3CE de calidad, innovación e investigación en docencia universitaria* (pp. 2580-2591). Alicante: Universidad de Alicante.
- Oficina Nacional de Normalización (1981). *Control de la calidad. Métodos de expertos*. NC 49:1981. Cuba.
- Ortega, F. (2008). El método Delphi, prospectiva en ciencias sociales a través del análisis de un caso práctico. *Revista Escuela de Administración de Negocios*, (64), 31-54. Recuperado de <https://journal.universidadean.edu.co/index.php/Revista/article/view/452/444>

- Piratoba, B. N. (2011). *Confiabilidad del instrumento para medir la habilidad de cuidado de cuidadores familiares de personas con enfermedad crónica en Cuidadores de personas mayores de la localidad de Usaquen, Bogotá, D.C.* (tesis de maestría). Universidad Nacional de Colombia, Colombia.
- Poblano, E., López, R., Gómez, J. and Torres, V. (2019). Effect of competitive intelligence on innovation capability: An exploratory study in Mexican companies. *Journal of Intelligence Studies in Business*, 9(3) 62-67. Doi: <https://doi.org/10.37380/jisib.v9i3.516>
- Quispe, A., Calla, K. M., Yangali, J. S., Rodríguez, J. L. y Pumacayo, I. I. (2019). *Estadística no paramétrica aplicada a la investigación científica con software SPSS, MINITAB Y EXCEL. Enfoque práctico*. EIDEC.
- Ramírez, J. S., Murcia, C. L., y Castro, V. (2014). Análisis de aceptación y preferencia del manjar blanco del valle. *Biotecnología en el Sector Agropecuario y Agroindustrial*, 12(1), 20-27. Recuperado de <http://www.scielo.org.co/pdf/bsaa/v12n1/v12n1a03.pdf>
- Ramírez, A. y Polack, A. M. (2020). Estadística inferencial. Elección de una prueba estadística no paramétrica en investigación científica. *Horizonte de la Ciencia*, 10(19), 191-208. Recuperado de <https://www.redalyc.org/journal/5709/570962992015/html/>
- Rodríguez, C. y Sossa, M. P. (2004). Validación de un cuestionario de conocimientos acerca de asma. *Revista Colombiana de Neumología*, 16(3), 162-168. Recuperado de <http://www.scielo.org.co/pdf/rcneum/v16n3/v16n3a2>
- Rodríguez, M. A., Poblano, E. R., Alvarado, L., González, A. y Rodríguez, M. I. (2021). Validación por juicio de expertos de un instrumento de evaluación para evidencias de aprendizaje conceptual. *RIDE Revista Iberoamericana para la Investigación y el Desarrollo Educativo*, 11(22). Doi: <https://doi.org/10.23913/ride.v11i22.960>
- Salas, R. S. (1998). *La evaluación en la educación superior contemporánea*. La Paz: Biblioteca de Medicina. Universidad Mayor de San Andrés.
- Van der, I. H., Goossens, L. H., Saatkamp, H. W. and Horst, S. H. (2002). Elicitation of quantitative data from a heterogeneous expert panel: formal process and application in animal health. *Risk Analysis*, 22(1), 67–81. Doi: <https://doi.org/10.1111/0272-4332.t01-1-00007>
- Vildózola, H. (2003). Organización del trabajo editorial [Presentación de trabajo]. XVII CCI FELSOCEM. Lima- Perú.

Wright, G., Lawrence, M. J. and Collopy, F. (1996). The role and validity of judgment in forecasting. *International Journal of Forecasting*, 12(1), 1-8. Doi: [https://doi.org/10.1016/0169-2070\(96\)00674-7](https://doi.org/10.1016/0169-2070(96)00674-7)

Rol de contribución	Autor (es)
Conceptualización	Rigoberto Reyes Valenzuela (principal), Domingo Noé Marrón Ramos (igual).
Metodología	Rigoberto Reyes Valenzuela (principal), Arturo González Torres (que apoya).
Software	Arturo González Torres (principal), Fátima Yaraset Mendoza Montero (que apoya), Rubén Juárez Rodríguez (que apoya).
Validación	Domingo Noé Marrón Ramos (principal), Rubén Juárez Rodríguez (que apoya), Rigoberto Reyes Valenzuela (que apoya).
Análisis formal	Rigoberto Reyes Valenzuela (principal), Arturo González Torres (igual).
Investigación	Arturo González Torres (principal), Rigoberto Reyes Valenzuela (igual), Fátima Yaraset Mendoza Montero (que apoya).
Recursos	Rigoberto Reyes Valenzuela (principal), Rubén Juárez Rodríguez (que apoya).
Curación de datos	Rigoberto Reyes Valenzuela (principal), Domingo Noé Marrón Ramos (igual), Fátima Yaraset Mendoza Montero (que apoya).
Escritura - preparación del borrador original	Fátima Yaraset Mendoza Montero (principal), Arturo González Torres (igual), Rubén Juárez Rodríguez (que apoya).
Escritura - revisión y edición	Rigoberto Reyes Valenzuela (principal), Domingo Noé Marrón Ramos (igual). Arturo González Torres (que apoya).
Visualización	Rigoberto Reyes Valenzuela (principal), Arturo González Torres (igual).
Supervisión	Domingo Noé Marrón Ramos (principal), Rigoberto Reyes Valenzuela (igual), Rubén Juárez Rodríguez (que apoya).
Administración de proyectos	Rigoberto Reyes Valenzuela (principal), Arturo González Torres (igual), Fátima Yaraset Mendoza Montero (que apoya).
Adquisición de fondos	Rigoberto Reyes Valenzuela (principal), Rubén Juárez Rodríguez (igual).